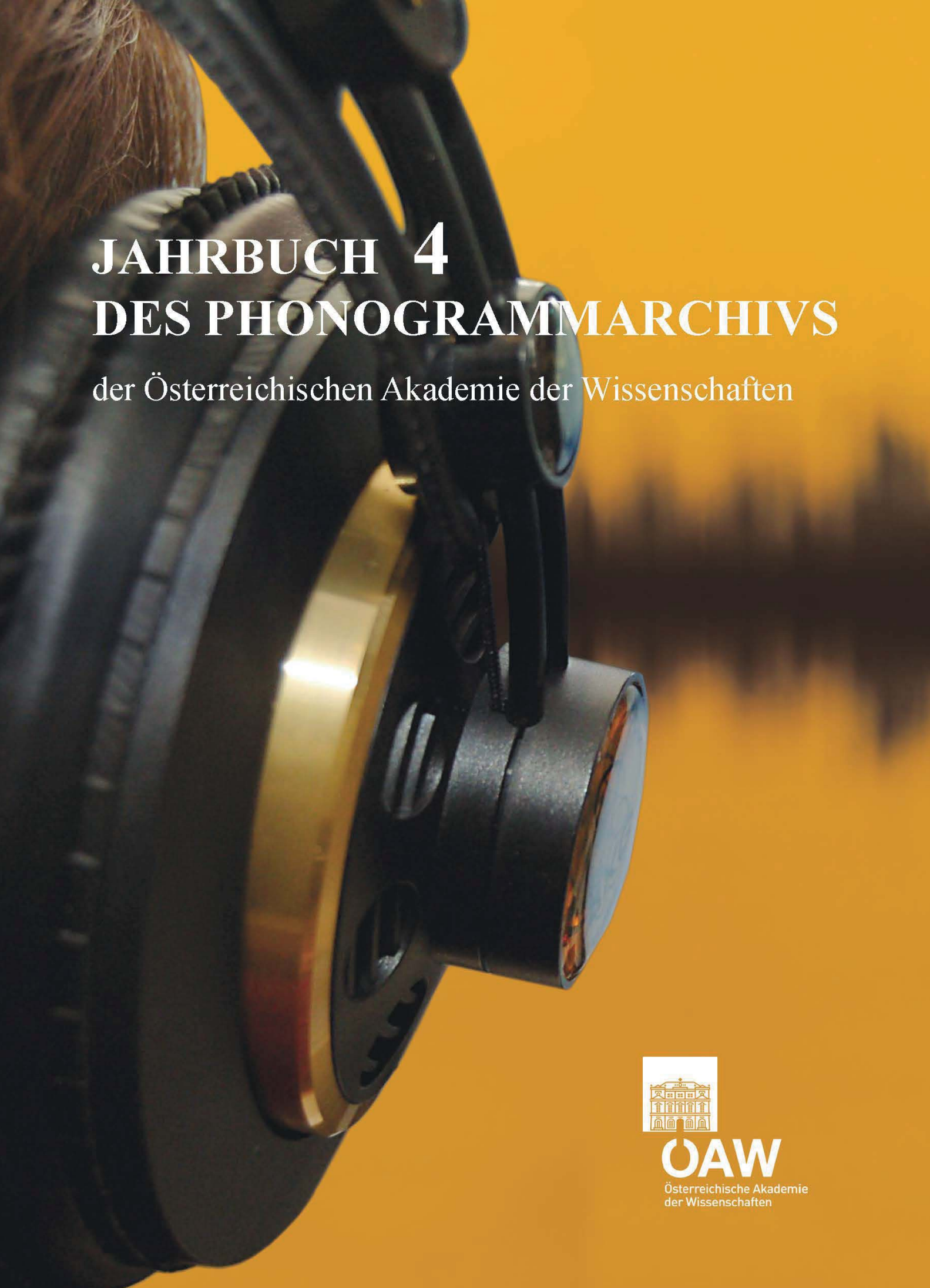


JAHRBUCH 4 DES PHONOGRAMMARCHIVS

der Österreichischen Akademie der Wissenschaften

ÖAW
Österreichische Akademie
der Wissenschaften

Jahrbuch des Phonogrammarchivs 4

Jahrbuch des Phonogrammarchivs

der Österreichischen Akademie der Wissenschaften

4

Gastherausgeber

Jürgen Schöpf

Herausgegeben von

Gerda Lechleitner & Christian Liebl

Cuvillier Verlag

Göttingen 2014

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2014

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2014
Gedruckt auf umweltfreundlichem, säurefreiem Papier
aus nachhaltiger Forstwirtschaft.

ISBN 978-3-95404-807-6
eISBN 978-3-7369-4807-5

Inhaltsverzeichnis

Vorwort

Dieser Band beinhaltet vornehmlich Beiträge zum Thema „Relationships of Speech Tone and Music“. Dem Prinzip der Kooperation zwischen dem Phonogrammarchiv und anderen Instituten der Österreichischen Akademie der Wissenschaften und der multidisziplinären Ausrichtung des Phonogrammarchivs folgend veranstaltete Jürgen Schöpf, Mitarbeiter am Phonogrammarchiv, gemeinsam mit dem Institut für Corpuslinguistik und Texttechnologie einen Workshop zu Tonsprachen und Musik. Die Präsentationen und Ergebnisse dieser wissenschaftlichen Veranstaltung sind nun in diesem Band des Jahrbuches nachzulesen. Das gewählte Thema, die daraus resultierenden Überlegungen und Methoden passen vorzüglich in die Arbeitsgebiete des Phonogrammarchivs, handelt es sich doch um ein relativ neues, bisher wenig beachtetes Forschungsgebiet, das Linguisten und Ethnomusikologen gleichermaßen anspricht. Dabei spielen verschiedene Aspekte, die auch den Arbeitsschwerpunkten des Phonogrammarchivs entsprechen, nämlich bedrohte Kulturen, Kontextualisierung und technische Spezialaufgaben, im Workshop-Thema „Beziehungen von Sprachton und Musik“ eine Rolle. Forschungen zu diesem Thema sind ohne Tonaufnahmen nicht möglich, manche von ihnen gehören zu den Beständen des Phonogrammarchivs. Zur Genese des Workshops, dessen Inhalten und Zielen ist in der Einführung von Jürgen Schöpf (s. S. 7–12) Genaueres zu erfahren.

Neben diesen 12 Beiträgen, die eine Keynote sowie Ausführungen zu Fallstudien aus Afrika, Asien und Grenzgebieten enthalten, gibt der Feldforschungsbericht von Barbara Kazianka einen guten Einblick in das von ihr gewählte Forschungsdesign, ihre Arbeitsmethoden und den (durch die jeweilige Situation „Aufnahmen im Kontext“ auch manchmal eingeschränkten) Einsatz von Audio- und Videotechnik. Kaziankas Forschungsinteresse gilt den Verflechtungen von Medizin und Identität bei den Wayuu, einer indigenen Gruppe auf der Halbinsel La Guajira (im Norden Kolumbiens und Venezuelas).

Im Blickwinkel des Phonogrammarchivs sind nicht nur neue Forschungsfragen, wie hier die Beziehungen von Sprachton und Musik oder die Sammlungszuwächse, dargestellt an einem Feldforschungsbericht, sondern auch einschlägige Publikationen, die aus der Beschäftigung mit Tönen als Quellen für spezifische Fragestellungen hervorgehen. Deshalb enthält dieser Band zwei Rezensionen von Veröffentlichungen, die einerseits auf historischen Tonaufnahmen (auch solchen aus dem Phonogrammarchiv) fußen, und andererseits frühe Tonaufnahmen, ausführlich kommentiert hinsichtlich der Entstehung und damaligen Zielsetzung, beinhalten. Maria Six-Hohenbalken erläutert umsichtig und reflexiv *Sensible Sammlungen*, eine kritische

Auseinandersetzung von Margit Berner, Anette Hoffmann und Britta Lange mit der akademischen Sammlungstätigkeit (v.a. in der Zeit bis zum Zweiten Weltkrieg), die einen neuen Umgang der Wissenschaftsgeschichte verschiedener Disziplinen, wie Anthropologie, Ethnologie, Ethnomusikologie etc., aufzeigt. Michaela Brodl stellt Jarmila Procházkovás Multimedia-Publikation (Kommentarband, Transkriptionen, Audio-CDs und Daten-CD) *As Recorded by the Phonograph* vor und unterstreicht die Bedeutung solcher Quelleneditionen. Beide Publikationen sind dem Aufgabenbereich „Kontextualisierung" zuzuordnen und geben ein wichtiges Feedback beziehungsweise eine Bestätigung zu ähnlichen Überlegungen, Diskussionen und Arbeiten im Phonogrammarchiv.

Auch wenn Band 4 des Jahrbuchs, ähnlich dem ersten Band, der Veröffentlichung von Beiträgen aus einer vom Phonogrammarchiv veranstalteten wissenschaftlichen Tagung geschuldet ist, bilden doch alle Beiträge zusammen – zwar regional und inhaltlich weit gefächert, aber in einer Trias zusammengefasst – die Archiv-Schwerpunkte „Bedrohte Kulturen", „Kontextualisierung" und „Technik" ab und werden durch den Feldforschungsbericht sowie die Rezensionen diesbezüglich ergänzt.

Gerda Lechleitner & Christian Liebl

Beiträge des internationalen Workshops „Relationships of Speech Tone and Music“

Introduction (Jürgen Schöpf)

Background

Relationships of speech tone and music have been intriguing me since my early studies in ethnomusicology and linguistics. Conducting fieldwork in Botswana in 1997 brought me in contact with a Bantu tone language of two tones, but in those days I had neither a matching methodology nor linguistic records that would have allowed me to focus on their relationship to singing and music – a recurring situation that is mentioned, among others, in Catherine Ingram’s contribution. The chance to participate in a DoBeS project in Upper Assam, India, documenting endangered Tai and Tibeto-Burman languages (isolating tone languages with contour tone) with project leader linguist Stephen Morey was a welcome challenge that taught me about the spectrum of possible relationships between speech tone and musics across different cultures. The “Workshop Relationships of Speech Tone and Music” in Vienna in 2012 developed out of the discourse between Stephen Morey and myself, and has become an important mile stone in what I see as an interdisciplinary basic research endeavour. Volumes dealing with this topic in a comparative way have previously been published, for example Thomas A. Sebeok & Donna Jean Umiker-Sebeok (1976) or Bonnie Wade (1993), but none of them targeted it as directly as the present one. The most valuable bibliography on the relation of speech tone and music is currently maintained online by Murray Schellenberg (2013); I had the pleasure of contributing a few items irregularly in the past.

Workshop “Relationships of Speech Tone and Music” (WRSTM)

The “Workshop Relationships of Speech Tone and Music” aimed at scholars from both linguistics and (ethno)musicology to contribute their methodologies, results and theoretical thinking. We had made it a point to provide travel money for students to attend. But in fact only one person in a pre-doc career status applied. In hindsight I believe that this was due to the relationship of speech and music being indeed a very demanding subject, requiring a certain level of expertise in both fields that is usually only developed in post-doctoral careers. The workshop from July 5th to 7th 2012 was rather small with two keynotes and a dozen papers, and an audience of thirty. But this manageable

size contributed to develop a laboratory atmosphere that allowed for an intensive exchange and mutual learning. At the end, I felt that all participants had profited from the workshop – and I would be pleased if this could spread through the present volume.

Review process

Each paper has been reviewed by both a musicologist and a linguist. Linguists were chosen by Stephen Morey, musicologists by myself. The pivotal criterion we had imposed on us was their familiarity with the language or area in question. Given that some contributions are dealing with very small minority languages it will be understood that not for all languages/areas complete matches were available. It is also obvious that in a small field full anonymity is impossible.

It is indeed a challenge working across the two disciplines, musicology and linguistics. Even basic terms like "tone", "pitch", "syllable", "rhyme", "intonation" and more have their specific discourses in both disciplines and they only rarely, or with long delays, inform each other (a major reason to hold the WRSTM). It is also unavoidable that a researcher's methodology is much more reflected, explicit and mature in his or her own discipline than in the respective neighbouring one. We have asked the reviewers to take this into account when reviewing the papers; yet it will always remain disputable what a useful balance should look like in that respect. In one case, a musicologist reviewer clearly recommended not to publish the paper of a linguist, for methodological reasons. We have, however, decided to include this contribution against this recommendation because we value it as a useful standpoint in the discourse with the other papers.

For lack of more compelling criteria I have arranged the papers geographically, starting in the West of Africa and ending in Japan, preceded by Bob Ladd's wonderful keynote. This keynote explains a number of phenomena in the music-speech relationship and their interplay, such as pitch, tone, intonation, contrast of tone and intonation, of grammatical versus lexical tone, and the important notion that tone in speech is not pitch alone, but includes other timbral parameters (very strikingly exemplified later on by Morey). Ladd's twofold concept of "top-down cues" and "phonetic residue" indeed provides a helpful tool to approach any tone language music relationship. A key to its understanding is that both provide redundancies in different ways to keep up intelligibility even though certain features of language are not represented in the singing. Moreover, I see the principles known from Gestalttheorie to be at work here for maintaining meaningfulness, such as the principle of "proximity", "similarity", and "common fate". It might be appropriate to add that music also comes with constraints that project onto language, e.g. by concepts of scales, an idea I have not yet seen reflected in linguistic studies.

“Mind the gap”

An interesting point becomes visible when arranging the articles geographically: the larger areas where tone languages occur (Africa, South East Asia, Middle America) appear mutually exclusive with areas where microtonal scale systems have developed (the Islamic world between Morocco and India). Although this is a very superficial observation I believe that tone languages profit from pitch contrasts beyond microtonic intervals to function efficiently (compare Vidal, quoted after Villepastour in this volume, saying “that spoken Yoruba would not normally place neighbouring speech tones less than a whole-step apart”). Whereas seen from the music end, a concept of musical scale that provides ample opportunities for microtonic expression in performance (systematically as a scale, not as ornaments) may not easily use the same perceptory parameter of pitch to develop the categorical (binary) contrasts needed to disambiguate meaning in language.

Amanda **Villepastour** has recently (2010) published an impressive work on Bàtá drumming, and in her present paper discusses methodological questions that I found particularly interesting, especially how the discourse in Yorùbá interacts with researchers. The concept of “do-re-mi-consciousness” is indeed an inspiration and calls for a subtle critique of our sources in this complex topic. She also reminds me that a plausible explanation is not necessarily a true one. Also, the often quoted continuum of speech over chant to song (quoting List, Bright and Agawu, see for example Ingram in this volume) is challenged by her data rather suggesting a continuum between speech and song, the latter exaggerating language features, whereas chant represents a performance practice less connected to both others. **Schellenberg,** at the far (eastern) end, has also recently been active in this area of research – see *Ethnomusicology* 56/2 (2012) – with a thesis in sharp contrast to Villepastour’s, which triggered very useful discussions at the WRSTM, to the end – as I see it – that the cultural relative as a research paradigm in ethnomusicology can make it difficult to argue with linguistics that is used to working with much more universal claims in its paradigm.

David **Locke**’s long research experience in West Africa allows for a broader and insightful overview more in qualitative terms than in detailed examples, including an interesting literature overview on the topic of West African music traditions. Concepts of the music of language (=drumming) and the language of music (= vocables) beautifully complement each other. That his analysis of drumming reveals three tone levels whereas the phonetics of the language distinguishes two should be of interest to linguists. Vanna **Crupi**, working in Uganda, brings us back to detailed analysis and field experiments showing current analytical methods as well as their limitations. Connecting “downdrift” and “descending melodies”, two old concepts of

linguistics and musicology respectively, with a conceptual background of neuroscience may be as instructive about the history of disciplinary gaps as it helps to understand her research data.

Jumping the (microtonal?) gap to Asia, **Morey** reveals detailed conceptions of a tone system that includes at least as many timbral qualities as actual pitches, and brings the historical depth of tonal changes into view. Other important lessons which I myself, as a musicologist, learned from him and in the Tai Phake community are that their speech tone system in its application to singing is hierarchical, resulting in the observable fact that syllables in rhyming position are treated differently than in non-rhyming position. In addition, a representation of contour tones appears more important than that of level tones. And not to forget: all these factors can lead to very different results for different genres within the same culture. **Schwörer**'s rich field work experience in south-east Asia allows to present an example of a very close relationship of text and instrumental music, including the perhaps surprising fact that this surrogate speech is not comprehensible to all members of the speech community. This finding is supported by **Ingram**'s data. Her analysis reveals the huge task that emerges once a language is not well documented, and, indeed, the methodologies available for researching speech tone music relationships are scarce. In her case, the pedagogical process is a key to access this relationship for Kam, which I believe is an important contribution to the methodology.

The collaborative work of **Karlsson, Lundström, Svantesson and Tuttle** is, in small, what the whole volume attempts: searching interdisciplinary methods and issues in the speech tone music relationships, and bridging linguistics and musicology. A tonal and a non-tonal dialect of Kammu are analysed in contrast, and Siri Tuttle, an authority on Athabascan, brings in her expertise on this native American language, the only example from the Americas in this volume. A problem that occurs across tone language cultures is how to separate tone from intonation, a problem already addressed by Svantesson in the past.

Jähnichen digs through amounts of data revealing a number of mechanisms in place between language and music in Vietnam, and is repeating Schwörer's find of surrogating speech of a musical instrument, this time a lute.

Schellenberg explains his experimental methodology in more detail than in previous publications, contrasting Mandarin and Cantonese in laboratory experiments, using specifically composed songs, and on the way explaining a couple of linguistic concepts such as intrinsic pitch of vowels, influences of consonants in front of vowels, and intrinsic vowel duration. The journey is completed by **Fujita**, working on Japanese where the rhythmic

component in language is of major interest. This parameter, although not directly discussed during the WRSTM, of course plays an important role, which is also touched on by Karlsson, Lundström, Svantesson and Tuttle.

Recurring issues

Overall, a number of recurring issues emerge. In my reading it is obvious that it is still a big methodological challenge to research the speech tone music relationship. This is explicit or implicit in most contributions, and due to different reasons. A major and structural one is the disciplinary gap between musicology and linguistics (I imagine that this gap would not exist had the humanities been developed by native speakers of Yoruba or Vietnamese). The theoretical paradigms of more universalist tendencies in phonetics and phonology on the one hand (claims like the "International Phonetic Alphabet", or the simplifying assumption that all speakers of a language share the same knowledge of that language), and a rather cultural relativist paradigm in ethnomusicology on the other, take their toll, making it difficult to understand each other over this disciplinary gap. A second challenge is that linguistic data on tone systems of languages is still scarce or difficult to access for musicologists, and sometimes it is even controversial; and a third one is that linguists have not yet discovered musicological studies as sources of information that can help them support their findings, e.g. how scale definitions and tone systems might connect. As long as only one side sees the opportunity it is difficult to develop cooperative project proposals.

In the present volume, I believe, the enthusiastic laboratory atmosphere from the WRSTM has matured into a batch of contributions that collectively improve our methodologies considerably. I hope that this topic will become increasingly addressed in both disciplines as methods advance and get exchanged more frequently. I would be pleased to be able to further this through present and future work.

Acknowledgements

A big word of thanks is appropriate to a number of institutions and people: the Phonogrammarchiv and the Institute for Corpus Linguistics and Text Technology, both part of the Austrian Academy of Sciences; the VolkswagenStiftung, and the Österreichische Forschungsgemeinschaft, who all contributed infrastructure and finances to the Workshop Relationships of Speech Tone and Music. Thanks to (now former) Präsident Helmut Denk of the Austrian Academy of Sciences, who inaugurated the WRSTM, and to the staff of the Phonogrammarchiv for their assistance during the workshop; a special thanks to former director Rudolf Brandl for supporting my

interest in this subject early on, and my personal congratulations on his 70th anniversary; thanks also to his successor, current director Helmut Kowar. Big thanks to Stephen Morey, who from the beginning developed the idea of the workshop with me and assisted with his experience in designing and running an appropriate review process. Thanks to my co-editors Gerda Lechleitner and Christian Liebl, especially for their patience in an endeavour continuously exceeding every deadline. Thanks to Elke Salzer for the layout. And, of course, to all authors (listed at the end of the volume) and the reviewers listed below:

Wolfgang Bender, Center for World Music, Hildesheim, Germany
Roger Blench, Kay Williamson Educational Foundation, Cambridge, UK
Peter Cooke, Professor Emeritus, University of Birmingham, UK
Jerold Edmondson, University of Texas at Arlington, USA
Catherine Falk, Melbourne Conservatory of Music, Australia
Ingrid Fritsch, Universität zu Köln, Germany
Nerida Jarkey, Sydney University, Australia
Randy La Polla, Nanyang University, Singapore
Debbie Loakes, University of Melbourne, Australia
Andreas Meyer, Folkwang Hochschule, Essen, Germany
Terry E. Miller, Professor Emeritus, Kent State University, Ohio, USA
Martin Mous, Leiden University, The Netherlands
Knut Olawsky, James Cook University, Townsville, Australia
Sophie Salffner, SOAS, London, UK
Jürgen Schöpf, Phonogrammarchiv, Austria
Paul Sidwell, Australian National University, Canberra, Australia
Scott Swan, Florida State University, USA
Marija Tabain, La Trobe University, Melbourne, Australia
Mike Tian Qiao Lu, James Cook University, Townsville, Australia
Trang Quang Hai, CNRS, Musée de l'Homme, Paris, France
Bell Yung, Professor Emeritus of Music, University of Pittsburgh, Pennsylvania, USA

References

Schellenberg, Murray. 2013. "Singing in Tone Languages - Partial Bibliography". Online: <http://www.linguistics.ubc.ca/singing_tones_biblio> (20/12/2013).

Sebeok, Thomas A. & Donna Jean Umiker-Sebeok. 1976. *Speech Surrogates: Drum and Whistle Systems*. 2 vols. The Hague & Paris: Mouton.

Wade, Bonnie C. 1993. *Text, Tone, and Tune*. New Delhi etc.: Oxford & IBH Publ.

2. Keynote

Singing in tone languages: An introduction to the kinds of things we might expect to find

(D. Robert Ladd)

This paper has two main goals. First, it gives a brief introduction to tone languages, aimed in particular at ethnomusicologists who are aware of the existence of such languages but are uncertain of exactly what it means to describe a language as tonal. Second, it sketches the kinds of interactions we might expect to find between musical pitch and linguistic pitch when people make music in a tone language. In addressing this second goal I draw extensively on our ongoing study of song in Dinka, a major language of South Sudan (<http://www.lel.ed.ac.uk/nilotic>).

The use of pitch in language

All spoken language, unless it is whispered, has pitch. Roughly speaking, the main physical basis of what we perceive as pitch is the fundamental frequency of regular vibrations, and it is physically impossible to have voice (vibrations of the vocal cords) without having a fundamental frequency (the rate at which the vocal cords are vibrating). In speech, the vocal cords vibrate at some (usually rapidly varying) frequency, and at the same time the tongue and lips and other articulators make rapid movements to produce vowels and consonants. The two systems operate fairly independently, so that a speaker can produce pretty much any syllable with any pitch and vice versa. This gives us some basis for thinking about the stream of vowels and consonants as being something separate from the pitch of the voice, and to some extent that separateness applies in all languages. For example, a voice raised in anger (where, among other things, the high pitches are higher) sounds very similar in any language.

However, all languages also use pitch for specifically linguistic purposes as well. For example, many languages distinguish questions from statements through pitch alone. The Italian sentence

(1) *Maria non viene* 'Maria not come-(3rd-person-singular)'

can be either a question meaning 'Isn't Maria coming?' or a statement meaning 'Maria isn't coming', depending almost exclusively on the pitch contour that accompanies the words. Such uses of pitch at the sentence level are usually called **intonation** (obviously, this term means different things in linguistics and in music). It is important to emphasise that intonation is *not* like the voice raised in

anger – intonation definitely varies from one language to another, although there are certain widespread properties. For example, consider the Hungarian sentence

(2) *József mérnök*, lit. 'Joseph engineer'

If it is spoken with a rise in pitch on *József* and a fall in pitch on *mérnök*, it means 'Is *Joseph* an engineer?' – that is, it is a question with the emphasis on *József*. To speakers of English or German, however, it sounds like a statement with the emphasis on *mérnök*, something like 'Joseph is an *engineer*'. Unless you know Hungarian, it is difficult to know which combination of emphasis and questioning or asserting is being expressed by the intonation here. For a general introduction to intonation see e.g. Cruttenden (1997) or Ladd (2008: chapters 1 and 2).

It is fairly common to categorise languages according to whether they use pitch at the sentence level only – as in the Italian and Hungarian examples we have just looked at – or also at the word level. The best-known example of a language that uses pitch at the word level is Chinese. Consider the following pair of Chinese[1] words:

(3) a. *tāng* 'soup' [moderately high level pitch]
b. *táng* 'sugar, sweets' [mid-to-high rising pitch]

Phonetically the difference between *tāng* and *táng* is just a matter of pitch, like the difference between a question and a statement in Italian or Hungarian. But functionally, it is obviously very different: in Chinese, the difference in pitch changes the meaning of the word, not just the general force of the sentence as a whole.

Pitch used at the word level is generally called **tone**, and languages that use pitch at the word level are generally called tone languages. Broadly speaking, there are three major areas of the world where tone languages are found: East and Southeast Asia, including New Guinea but not Australia; pretty much all of Sub-Saharan Africa; and various parts of the Americas, especially southern Mexico and the Amazon. Something like half the languages in the world are tone languages, and something like half the people in the world speak a tone language, so this is hardly a marginal phenomenon. But since the very idea of such a language strikes most speakers of European languages as outlandish, it is important to spend some time explaining a bit more about how they work. Further detail can be found in such works as Gussenhoven (2004) and Yip (2002).

1 Here and throughout the paper the examples from 'Chinese' are drawn from the standard language, often called Mandarin.

Basic structural properties of tone languages

It is surprisingly difficult to come up with a definition of a tone language that commands general agreement. On the other hand, it is not hard to identify a set of languages that everyone will agree on classifying as tone languages, a set of languages that everyone will agree are non-tonal, plus a set of languages that, for one reason or another, fall somewhere on a scale between definitely tonal and definitely non-tonal. I begin by describing the properties of languages that are definitely tonal.

The most important point is that pitch in tone languages is not some kind of overlay or musical supplement, but *an integral part of the word*. In order to talk about word forms in a tone language, it is not enough just to talk about the consonants and vowels; it is necessary to talk about the pitch and sometimes the voice quality as well. Let us go back to the Chinese example in (3). For a speaker of a European language, it is very tempting to talk about 'the word *tang*' and say that if you pronounce it one way it means 'soup' and if you pronounce it another way it means 'sugar'. But that is similar to looking at a pair like English *sweets* and *tweets* and concluding that the way you pronounce the first consonant makes a difference to the meaning of the word. Most English speakers would say 'No, *sweets* and *tweets* are different words', and that is exactly what Chinese speakers would say about *tāng* and *táng*. It is important to get used to the idea that in tone languages, pitch can work *just like* consonants and vowels for distinguishing one word from another.[2]

Tonal distinctions can signal various kinds of meaning differences. The most basic kind of difference is the **lexical** difference just illustrated: *tāng* and *táng* have a different tone and are different words. But tone can also make a **grammatical** difference. In many languages tonal differences are used to signal the difference between singular and plural, or between present and past, or other such grammatical distinctions. In Dinka, for example, we get singular-plural pairs like the following:[3]

(4) a. *kët* [L] 'shoulder'; *kët* [H] 'shoulders'
b. *joth* [L] 'ring'; *joth* [H] 'rings'

On the basis of these two cases only, it might seem that we can talk about 'the word *kët*' or 'the word *joth*' and treat the pitch difference – low in the singular, high in the plural – as something added. But if we look at Dinka grammar as

2 It is not just pitch that can work this way; in some languages voice quality can have the same kind of effect. In Dinka, *roor* (with modal or creaky voice) means 'forest' and *röör* (with breathy voice, indicated by the dieresis on the vowel letter) means 'men'. However, in this paper I concentrate on pitch.

3 The examples are given in Dinka orthography, which does not indicate tone. Tone is specified here as low [L] or high [H]. The dieresis on a vowel indicates breathy voice quality; cf. the previous footnote.

a whole, we see that the change in these forms is more like the difference in English between *woman* and *women* or *mouse* and *mice* – in both cases what we are doing is *modifying* the word, and the *modification itself* is the signal of the grammatical category. Plural marking in Dinka is almost always a question of changing some part of the word, and it can be almost any part, not just the tone. This can be seen in these further examples:

(4) c. *dhɔk* [L] 'boy'; *dhäk* [L] 'boys'
d. *meth* [L] 'child'; *mïth* [L] 'children'
e. *nɔk* [H] 'feather'; *nak* [L] 'feathers'
f. *jïc* [H] 'ear'; *jïth* [L] 'ears'

Here we see: pairs where the vowel is different, like (4c); pairs where the vowel and the voice quality are different, like (4d); pairs where both the vowel and the tone are different, like (4e); and pairs where the final consonant and the tone are different, like (4f). Once again, in short, tone is behaving exactly like vowels and consonants.

Now let us consider the kinds of **phonetic** distinctions that are involved in tone. A very common traditional distinction is the difference between **level** tones and **contour** tones. In the case of level tones, the idea is that the pitch on each syllable is relatively unchanging, so that the linguistic distinctions are between categories like high and low – as in the Dinka examples just given. With contour tones, the distinctions are between categories like rise and fall. As a very broad generalisation, the tone languages of East and Southeast Asia involve contour tones, whereas the languages of sub-Saharan Africa involve level tones. As a result of this broad typological difference, the descriptive and notational conventions of linguists and anthropologists working in those two areas are also rather different.

A brief illustration will suffice. Here are the tones of Chinese, along with the standard tone diacritics.

(5) a. *fū* [high level] 'hatch [verb]'
b. *fú* [mid-high rising] 'submit'
c. *fǔ* [low dipping-rising] 'rot, ferment'
d. *fù* [high-low falling] 'pair, set'

What can be seen here is that the diacritics are a rough visual indication of the pitch pattern across the syllable. This is fairly typical of notation schemes for the contour tones of the tone languages of East and Southeast Asia. It is also fairly typical that among the contours there are one or more levels; the Chinese tonal inventory, for example, includes a high level tone in addition to a rise and a fall and a sort of low dip. The Chinese examples can be compared with a lexical set from Yoruba, a major language of Nigeria.

(6) a. *īgbā* [mid – mid] ‘200’
b. *īgbá* [mid – high] ‘calabash’
c. *ìgbà* [low – low] ‘time’
d. *ìgbá* [low – high] ‘[type of tree]’

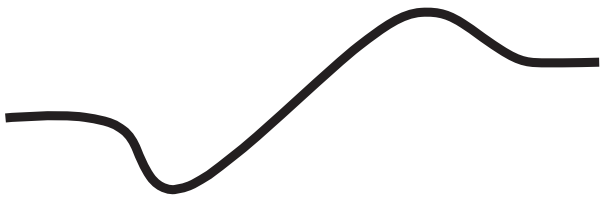
In writing and transcribing African languages the acute and grave accents tend to be used to indicate high and low, and if a language also has a mid tone, which Yoruba does, it may be marked like in the way illustrated in (6a) and (6b), or just left unmarked. Ethnomusicologists working in different parts of the world may need to be aware that these different descriptive conventions exist; tone marks may mean different things in descriptions of different languages.

It is also worth mentioning that in many African and American tone languages, especially in West Africa and Mexico, even though the system is essentially based on level tones, there are many cases where contour tones occur phonetically but count as a combination of two levels. For example, consider the following Yoruba sentence:

(7) a. Ayọ̀ lọ ‘Ayo [man’s name] is coming’

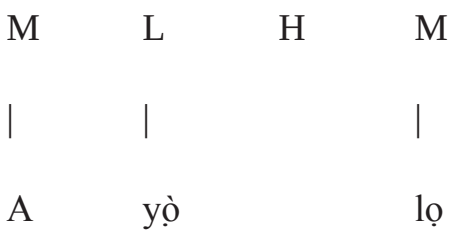
The name *Ayọ̀* has a sequence of a mid and a low syllable, while the verb *lọ* has a mid tone. However, the grammar requires a high tone that (very approximately) marks the end of a subject noun phrase, and this high tone attaches itself to the second syllable of *Ayọ̀*. This gives a low-high sequence on that syllable, which is pronounced as a steep rise before going back down to the mid-level for the verb *lọ*, so that the pitch contour on the whole utterance looks roughly like this:

(7) b.



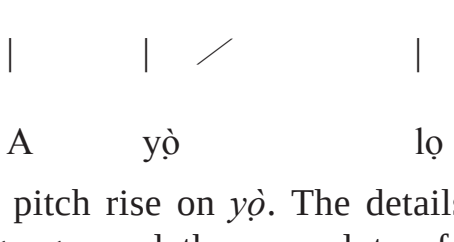
Tonal phenomena like this are the basis of so-called autosegmental notation that is often seen in linguistic descriptions. In an autosegmental analysis, the tones form one sequence of units and the syllables form another sequence. Some of the tones come preattached to syllables, and some do not; the tones that do not come preattached, like the H tone that marks the end of the subject in Yoruba, have to line themselves up according to various principles. That is, we start with a structure like this:

(7) c.



and then the 'post-lexical' attachment of the subject-final H tone gives us this structure:

(7) d.



which yields the pitch rise on *yọ̀*. The details vary to some extent from language to language, and there are lots of theories about the general structural principles that govern the association of tones to syllables. Unfortunately, I can do no more than mention the existence of autosegmental notation here; for further detail the reader may consult Goldsmith (1990).

'Pitch accent' and other problems

There are many other linguistic aspects of tone that could be discussed here – whole books have been written about tonal phonology – but the foregoing brief introduction will have to do. Before moving on to singing, however, it is important to mention why there are borderline cases – what it means for a language to be hard to classify between tonal and non-tonal. The basic source of the uncertainty is the existence of accent.[4]

Many languages, including English, pick out one syllable of a word or phrase for what I will refer to informally as 'special phonetic treatment' of some sort. This is a good rough-and-ready definition of accent. Just like tone, accent can depend on both lexical and grammatical factors. In Italian, for example, we have pairs like the following:

(8) a. ***an****cora* 'anchor'; *an****co****ra* 'again, still'
b. ***prin****cipi* 'princes'; *prin****ci****pi* 'principles'

Here the difference in accent placement (indicated by graphically emphasising the accented syllable) distinguishes two different words. There are also pairs like the following:

4 This section draws especially on ideas from Beckman (1986) and Hyman (2006, 2009); see also Himmelmann & Ladd (2008).

(9) a. ***par****lo* ‘I speak’; *par****lò*** ‘s/he spoke’
b. ***la****vati* ‘wash! [imperative];
*la****va****ti* ‘washed’ [masc. pl. past participle]

In these the different accent placement signals the grammatical difference between two forms of the same verb.[5]

Also, as with the difference between tone and intonation, accent can be used at the sentence level as well as at the word level. This is most easily seen in English:

(10) a. I **thought** he was coming.
b. I thought he was **coming.**
c. I thought **he** was coming.

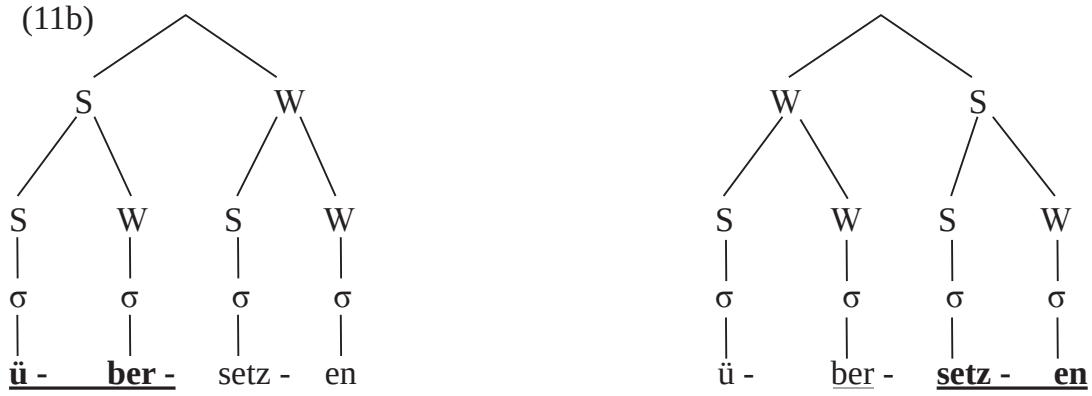
These three sentences involve the same words and in some sense have the same propositional meaning, but would be used in different circumstances: 10a would likely be used when the speaker has just had confirmation that ‘he is coming’, whereas 10b would likely be used when the speaker has just learned that this is not true after all, and 10c when the speaker learns that someone else is coming instead. Finally, in some languages accent is organised hierarchically, so that there are metrical structures that are reminiscent of music. This is very definitely true of English and German. In German for example, there is a distinction in pairs like the following:

(11a) ***über****setzen* ‘take across’; *über****setzen*** ‘translate’

At one level of analysis, this distinction involves the relative prominence of the prefix *über* and the stem *setzen*. But within both of those component parts – at the next lower level, as it were – there are also differences of relative prominence, with the first syllable in each case being more prominent than the second. This nesting of relative prominence can be represented graphically as a ‘metrical tree’, as follows (the symbol σ stands for ‘syllable’; *s* and *w* stand for ‘strong’ and ‘weak’ respectively):

5 For the sake of academic rigour, two things should be pointed out here. In (9)a, the diacritic on the final syllable of *parlò* is part of standard Italian orthography and indicates that the final syllable is accented; accent on other syllables is not normally indicated in spelling. In (9)b, the imperative form actually has a different internal structure from the past participle form, in that it is a sequence of the imperative *lava* and the second person reflexive pronoun *ti*; on the surface, however, the only difference is the placement of the accent.

(11b)



This kind of structure is central to accentual patterns in English and German, but it may be a relatively unusual phenomenon cross-linguistically. The basic metrical idea is due to Liberman (1979); for more on metrical phonology see e.g. Hayes (1995). In any case, the essence of accent is taking a syllable and singling it out for special phonetic treatment – or, if there is hierarchical structure like that shown in (11a), it is a matter of singling out one element in a larger group, at multiple levels. That much seems to be a valid cross-linguistic generalisation. What can vary a lot from one language to another, however, is the actual phonetic nature of the special treatment that signals which element is singled out. In Italian and English and German, the special treatment involves what is usually called **stress** – the accented syllable is pronounced with greater force of articulation, it is often longer in duration, and there are a variety of pitch phenomena connected with the way intonation patterns are associated with words. In other languages the special treatment may be different, and in particular, it may be primarily or exclusively a matter of pitch. Moreover, in languages where the accent is signalled mainly by pitch there is often no hierarchical organisation of the kind we just saw in English or German, but only a syllable-by-syllable specification: accented or unaccented. Languages like this are sometimes called 'pitch accent languages', but there is no general agreement on the boundary between 'pitch accent' and tone.

The problem is easily seen in Standard Japanese. The main phonetic manifestation of accent in Japanese is a drop in pitch from the accented syllable to the following syllable. The following is a standard example:

(12) a. ***ha****si desu* 'they're chopsticks'
b. *has**i** desu* 'it's a bridge'

In ***ha****si* 'chopsticks' the pitch drops from the first syllable to the second, whereas in *has**i*** 'bridge' the pitch drops from the second syllable to the first syllable of the next word. This still seems easy to distinguish from tone – and it still seems a lot like English or German or Italian, except for the

phonetic detail. However, the problem is that Japanese also has words *with no accent*. That is impossible in English or German or Italian – in any word of more than one syllable, one of the syllables has to be more prominent than the others. In a word with two syllables, like English *permit* or Italian *parlo/parlò*, there are only two possibilities – either the first syllable is accented, or the second syllable is accented. But in Japanese there is a third possibility:

(13) c. *hasi desu* 'it's an edge'

Phonetically, this is distinguished from (13) b. by the fact that there is no pitch drop during or after the unaccented word *hasi* 'edge'.

When we consider these two characteristics – the existence of unaccented words, and the fact that accent, when it is present, is manifested mostly by pitch – we suddenly see that Japanese seems rather different from English or German or Italian. There are quite a few languages that have accentual systems of this general sort. In these cases, some people emphasise the similarity to accentual systems as in English or German, and others emphasise the similarity to tone languages. There is definitely a scale or continuum, and this is why a given language may be classified both as 'tonal' and 'non-tonal' by different researchers. There is still no generally accepted theory of how to fit pitch accent systems into a valid typology.

Singing in a tone language

Now let us consider the kinds of things that we might expect to find when people make music in a tone language. People who speak non-tonal languages tend to assume that there is a fundamental problem here, but any problem is entirely analogous to other situations in which language is used in special ways. For example, when people ask me 'how do they sing in a tone language?', I always want to respond with something like 'how do you whisper in a language with voicing distinctions?'. In many languages, including almost all European languages, there are consonant distinctions based on the presence or absence of voicing. The difference between [s] and [z], for example, is a difference of whether the vocal cords are vibrating or not.[6] When we whisper, the vocal cords are not vibrating. Yet it is not hard to understand a whispered sentence like *Sue went to the zoo* or *There were fleas in the fleece.*

6 This difference is readily observed by pronouncing [s] and [z] alternately, while either (a) placing the thumb and forefinger on either side of the Adam's apple to feel the vibration, or (b) blocking the ears to hear the vibration directly through the bones of the neck and head.

There are at least two distinct aspects to what makes this possible. One is the existence of what are often called **top-down cues**. When we hear a sentence, we try to make everything fit together grammatically and lexically. *Sue went to the zoo* makes sense; *Zoo went to the Sue* does not. A sentence that begins *There were* expects a plural noun following, so *fleas* is more likely than *fleece*. This kind of feature figures prominently in song, where texts are often fairly predictable or traditional, and may also involve a lot of repetition.

But there is also a basic phonetic aspect to why it is possible to hear the difference between [z] and [s] even when voicing is switched off for whisper. This is the fact that there are a lot of *other* subtle phonetic differences besides voicing that contribute to our perception of the difference. One difference is that the friction in [s] is a lot more strident than the friction in [z]. Another is that, in English, vowels before voiced consonants are a lot longer than before voiceless consonants, so that the difference between *fleas* and *fleece* is signalled not only by the voicing of the consonant but also by the length of the vowel – which of course is not affected by whispering. These and other predictable accompaniments to the basic voicing distinction tend to remain even when the voicing itself is overridden, and this helps us perceive the distinction. I will refer to this kind of phenomenon as **phonetic residue**.[7]

Now let us consider how this applies to tone languages. Much of the answer to how it is possible to sing in a tone language is essentially the same as the answer to how it is possible to understand whispering in a language with voicing distinctions: first, there are lots of top-down cues, as there are in any language, and second, there are lots of predictable phonetic differences besides pitch that survive as phonetic residue and give us additional cues to the tonal categories. A standard example of these predictable phonetic differences comes from Chinese, where syllables with the four different tones (illustrated in (5) above) have quite different typical durations, with the falling tone the shortest, the low dipping tone the longest, and the other two in between (Howie 1976). Similarly, in many languages there are subtle differences of vowel quality when 'the same' vowel phoneme is pronounced with different tones.

Predictable phonetic differences of this sort are likely to be maintained as phonetic residue in singing. For example, we might find that the predictable durational differences between tones in Chinese can be observed in singing.

7 It is worth emphasising that the predictable accompaniments that contribute to phonetic residues may be based on universal phonetic tendencies – things that result from the mechanics of the speech production system – or they may be language-specific structural features. The case of vowel length before voiced and voiceless consonants is a bit of both: in many languages there is a tendency for vowels to be longer before voiced consonants, but the ***extent*** of the difference in English is extreme.

(I know of no research specifically on this question, but it is known – cf. Liu & Samuel 2004 – that these differences are preserved in whispering and that listeners make use of them in understanding whispered Chinese speech). Similarly, in languages where vowel quality is affected by different tones, we might expect to find those differences in sung vowels even though the pitch itself is determined by the music. We have found such effects in our Dinka study. Specifically, we have found that there are subtle differences in the precise phonetic quality of breathy vowels and non-breathy vowels in ordinary speech (technically, differences of vowel height), which are mirrored in singing even though the voice quality difference itself is overridden by the fact of singing (cf. Rognoni 2011).

Another kind of phonetic effect that could help signal tonal distinctions would be more direct: we might find that the realisation of the musical pitch directly reflects something of the linguistic pitch. For example, we might find that a linguistically rising tone shows a slight rise at the beginning or end of the musical note. We have looked for effects of this kind in Dinka, but have not found them; however, they are reported to occur in Cantonese opera (Yung 1989). Another conceivable effect of this sort would be that linguistically high tones are produced slightly sharp and linguistically low tones slightly flat. We looked for this in Dinka as well, but we have found no evidence at all that it happens.

In short, there are many subtle ways in which the phonetic identity of linguistic contrasts based on pitch could come through in singing. On the whole, though, the phonetic research needed to establish which of these ways are actually found has not been done, or has been done only in limited ways.

Text-setting constraints

There is a further interesting aspect to song and to poetry more generally, which also seems to play a significant role in making it possible to sing in a tone language. This is the existence of what I will refer to as **text-setting constraints** – rules about how a musical or poetic text fits the music or the verse.

In English and many other European languages, text-setting pays attention to accent**.** Poetic lines and musical phrases include strong and weak positions, and, in English at least, an unaccented syllable cannot appear in a strong position. (The reverse is not true: an accented syllable *can* appear in a weak position.) In poetry, strong positions are defined by the metre; there is a regular metrical pattern, and part of the art of writing poetry is to find words that fit it. On a very trivial level, consider the English limerick. The metrical pattern of the first line of a limerick is

(14) da DAH da da DAH da da DAH (da)

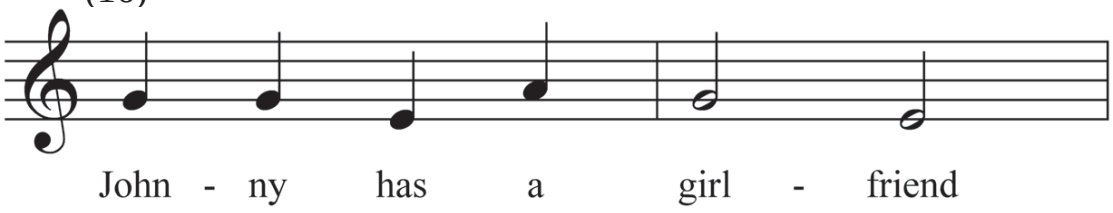
i.e. with three strong positions separated by two weak positions. It is very common for limericks to begin with a line introducing the main character and the place the main character comes from:

(15) a. There **once** was a **man** from Ma**dras**
b. There **once** was a **la**dy from **Ire**land
c. There **was** an old **man** from Stam**boul**

For most part the place is irrelevant, and tends to be chosen to rhyme with the rest of the limerick. But the metre does place constraints on who can come from where. If the main character is a 'lady', she cannot come from a place beginning with an unaccented syllable, because that will result in three weak syllables between the strong ones (**There* ***once*** *was a* ***la****dy from Stam****boul***). If the main character is an 'old man', the line must start with *There was* instead of *There once was* (**There* ***once*** *was an old* ***man*** *from Ma****dras***), for the same reason. And so on.

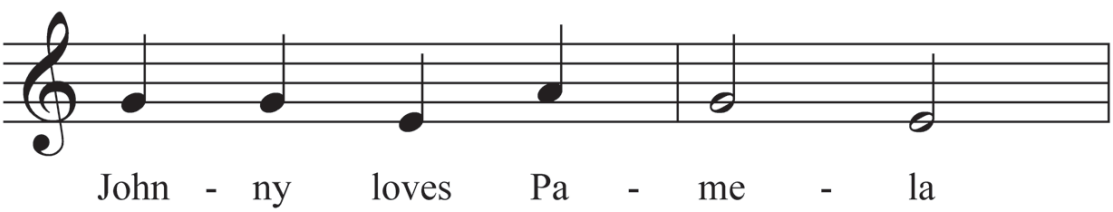
Similar things are true in music. In this case, 'strong position' basically refers to the musical downbeat. Every English-speaking five-year-old knows this constraint and how to work with it, because of the existence of a playground taunt with a simple melody:[8]

(16)



Here the two downbeat positions are occupied by accented syllables. But what happens if a child wants to draw attention to a playground romance and chant *Johnny loves Pamela*? There are six notes in the tune and six syllables in the sentence, but every five-year-old knows that it is impossible to sing

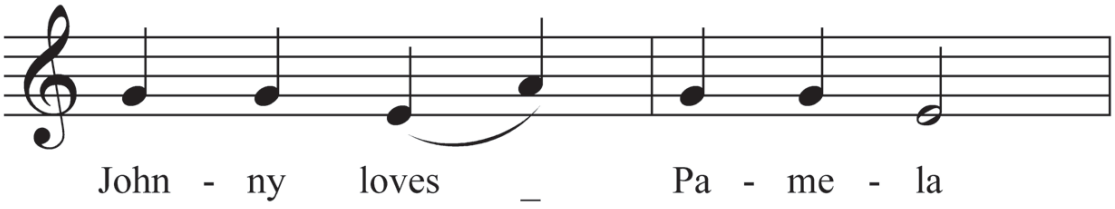
(17) a.



8 The specific examples here are from Ladd (2008: 57), but the linguistic significance of the children's taunting melody was first discussed by Liberman (1979).

Placing the unaccented syllable *-me-* on the beat makes the words all but unintelligible. Rather, it is necessary to rearrange things so that the *accented* syllable of *Pamela* comes on the downbeat; we create a melisma on *loves*, subdivide the note on the second downbeat, and get:

(17) b.



On a slightly more elevated level, the same principle is illustrated by the famous missetting of the word *incorruptible* in the aria *The trumpet shall sound* in Handel's *Messiah.* Handel set the word as if it were pronounced *incorrup**ti**ble*, which of course is something like what it would have been in his native German. But the result is almost unsingable for English native speakers, and the setting is usually edited to put the accented syllable *-rup-* on the downbeat (e.g. Shaw 1958: 190ff.).

Now let us return to tone languages. Just as the downbeat-to-accent constraint helps English singers and listeners, there is evidence that in tone languages there are textsetting constraints based on pitch that in some way help the singer to feel the melody, and help the listener to recover the underlying tonal pattern of the words in the text. The existence of such a constraint was first investigated experimentally in Cantonese by Wong & Diehl (2002). Essentially, the constraint is this: if the linguistically specified pitch direction across a pair of syllables is **up**, then the musical melody across the corresponding notes should go up; if the linguistically specified pitch direction across a pair of syllables is **down**, then the musical melody across the corresponding notes should go down; if the linguistically specified pitch direction across a pair of syllables is **level,** then the musical melody should either stay level or go up, but not go down.

This sounds very restrictive, but like the downbeat-to-accent constraint in English, it actually allows quite a few possibilities. First of all, it depends on a broad classification of the six tones of Cantonese into only three classes, which Wong & Diehl call high, mid and low. Consequently, the sequence 'high-high' covers four distinct possible tone sequences. Second, the constraint is concerned only with pitch direction, not the amount of pitch change: low-mid, low-high and mid-high all count as 'up', and, on the musical side, the melody can go up by different intervals that still all count as 'up'. Third, if the linguistic melody is level – that is, if the tone

stays the same on two successive syllables – it is compatible with both level and rising musical melodies. So there is quite a bit of latitude, but the framework defined by this broad constraint seems to be observed by both composers and listeners. There were astonishingly few violations of the constraints in the songs Wong & Diehl analysed, and in their experiments, listeners understood experimental song texts – which were designed to be ambiguous – as if they followed the rules. Subsequent work by Schellenberg (this volume) and others (e.g. Ho 2006, Chow 2011) has confirmed Wong & Diehl's basic finding.

In another study, Schellenberg (2009) looked to see if the same kind of constraint was operative in Shona (a Bantu language spoken in Zimbabwe). Shona is a very different language, where tone is much less essential for word recognition than it is in Cantonese, but Schellenberg found some evidence for a weaker version of the same constraint. Specifically, what he found was that if the linguistically specified pitch direction across a pair of syllables is up, then the musical melody across the corresponding notes should not go down; if the linguistically specified pitch direction across a pair of syllables is down, then the musical melody across the corresponding notes should not go up. If the linguistic melody is level – that is, if two consecutive syllables have the same tone –that seems to be compatible with any musical note sequence, and level musical melody seems to be compatible with any linguistic tonal sequence. His analysis is somewhat different from Wong & Diehl's, so it is difficult to make a direct comparison, but the same basic preference for the musical melody to match the linguistic melody seems to be present in both languages; at the same time, it also seems that the constraint is less restrictive, and that violations of the constraint are more common in Shona than in Cantonese. Table 1 shows a rough comparison of Wong & Diehl's findings with Schellenberg's.

	Linguistic tone sequence	Musical note sequence up	Musical note sequence down	Musical note sequence up
Cantonese	Up	Preferred	Rare	Absent
	Down	Rare	Preferred	Absent
	Level	OK	Absent	Preferred?
Shona	Up	Preferred	Dispreferred	OK?
	Down	Dispreferred	Preferred	OK?
	Level	OK?	OK?	Preferred?

Table 1: Text-setting constraints for Cantonese (Wong & Diehl 2002) and Shona (Schellenberg 2009).

In our Dinka project we have looked for this constraint as well. We had actually previously concluded that tone was simply ignored in singing, because we were unable to find any phonetic residue effects of the sort

discussed in section 2.1. With some idea of what to look for, however, we have applied statistical analyses to our song corpus similar to those used by Schellenberg, and we have tentative findings of a constraint very similar to what he found for Shona – a text-setting constraint based on pitch direction. The work is still in progress and it is too early to be sure of our final conclusion, but it does appear that pitch direction plays at least some role. My student Inga McKendry, who has worked for many years on the Mixtec language group in Oaxaca, Mexico, reports informally (personal communication, June 2012) that pitch direction seems to play a key role in setting texts to music in Mixtec as well.[9]

Conclusion

To summarise: I suggest that there are two fundamentally different kinds of answers to the question 'how can they sing in a tone language?', and at least two different kinds of effects that we should look for when investigating this question. First is the existence of residual phonetic cues to tonal contrasts, and second is the existence of text-setting constraints. I hope that future research can help establish whether this is a helpful dichotomy – and also whether these two kinds of effects cover everything, or whether there are still other kinds of manifestations of tone in singing. It will also be interesting to see whether it is just a coincidence that text-setting constraints seem to be more restrictive in Cantonese, where tone is often the only cue in context to a word's meaning, than in Shona or Dinka, where tone is much more predictable. Finally, it will be interesting to see whether some of the typological uncertainties regarding pitch accent in linguistics are reflected in any way in the type of text-setting constraints found in musical traditions in different languages.

9 It is worth noting in this connection that Shona, Dinka, Cantonese and Mixtec are about as far apart geographically and typologically as it is possible for languages to be. If all of these languages base the relation between linguistic tone and musical melody on pitch direction, that fact is definitely of interest for our understanding of pitch perception in both speech and music. It is well established (see the summary in Patel 2008, ch. 4) that pitch interval and pitch direction are in some sense processed separately in the brain. If preserving pitch direction turns out to be the most important thing in setting linguistic pitch to music, this has a lot of implications for possible theories of how pitch is processed, and ultimately for the larger question of which cognitive resources are shared by music and language and which are independent.

References

Beckman, Mary E. 1986. *Stress and non-stress accent.* Dordrecht: Foris Publications.

Chow, Man-Ying Sheryl. 2011. "Text and Melody in Cantonese Popular Music: A Relationship of Subordination or Complementarity?" *Proceedings of the Second Asian Conference on Arts and Humanities 2011:* 12–32. [Online: <http://iafor.org/acah_proceedings.html>]

Cruttenden, Alan. 1997. *Intonation.* 2nd ed. Cambridge etc.: Cambridge University Press.

Goldsmith, John. 1990. *Autosegmental and metrical phonology.* Oxford: Blackwell.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation.* Cambridge etc.: Cambridge University Press.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies.* Chicago etc.: University of Chicago Press.

Himmelmann, Nikolaus & D. Robert Ladd. 2008. "Prosodic description: An introduction for fieldworkers". *Language Documentation and Conservation* 2: 244–274.

Ho, Wing See Vincie. 2006. "The tone-melody interface of popular songs written in tone languages". In: Baroni, Mario et al. (eds.). *Proceedings of the 9th International Conference on Music Perception and Cognition.* Bologna: Alma Mater Studiorum University of Bologna, 1414–1422.

Howie, John Marshall. 1976. *Acoustical studies of Mandarin vowels and tones.* Cambridge etc.: Cambridge University Press.

Hyman, Larry M. 2006. "Word-prosodic typology". *Phonology* 23: 225–257.

Hyman, Larry M. 2009. "How (not) to do phonological typology: the case of pitch-accent". *Language Sciences* 31: 213–238.

Ladd, D. Robert. 2008. *Intonational phonology.* 2nd ed. Cambridge etc.: Cambridge University Press.

Liberman, Mark. 1979. *The intonational system of English.* New York: Garland.

Liu, Siyun & Arthur G. Samuel. 2004. "Perception of Mandarin lexical tones when F0 information is neutralized". *Language and Speech* 47: 109–138.

Patel, Aniruddh D. 2008. *Music, Language, and the Brain.* Oxford etc.: Oxford University Press.

Rognoni, Luca. 2011. "The neutralization of the voice quality distinction in Dinka songs: a production and perception study". In: Lee, W.-S. & E. Zee (eds.). *Proceedings of the 17th International Congress of Phonetic Sciences, 17–21 August 2011, Hong Kong*. Hongkong: City University, 1702–1705.

Schellenberg, Murray. 2009. "Singing in a Tone Language: Shona. In: Ojo, Akinloye & Lioba Moshi (eds.). *Selected Proceedings of the 39th Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 137–144.

Shaw, Watkins (ed.). 1958. [Handel's] *Messiah: vocal score*. London: Novello.

Wong, Patrick & Randy Diehl. 2002. "How can the lyrics of a song in a tone language be understood?" *Psychology of Music* 30: 202–209.

Yip, Moira. 2002. *Tone.* Cambridge etc.: Cambridge University Press.

Yung, Bell. 1989. *Cantonese opera: performance as creative process.* Cambridge etc.: Cambridge University Press.

3. Articles

Talking Tones and Singing Speech among the Yorùbá of Southwest Nigeria

(Amanda Villepastour)

In an article published in vol. 56 of *Ethnomusicology*, Murray Schellenberg (2012: 266) states, "language is not a determinant of music in tone languages, but rather […] music accommodates language when it is convenient but is perfectly willing and able to override linguistic requirements". He concludes, "Music is willing to accommodate language as long as this does not interfere with the music" (ibid.: 276).[1] While these statements may be true of many tone languages of the world, most particularly in Asia but also in some areas of Africa,[2] my research into the linguistic aspects of drumming and singing among the Yorùbá people of southwest Nigeria does not corroborate Schellenberg's opening and concluding statements, which imply that the relationship between speech tone and music is incidental rather than interdependent.[3] Rather, my data say that mimetic and coded Yorùbá drumming only rarely deviates from the pitch contours of natural speech. Likewise, chanted and sung text deviates from spoken and written Yorùbá in ways that are both formulaic and mimetic of natural speech's own "deviations" from expected speech tones, rather than serving the aesthetic needs of music, as many scholars have repeatedly claimed. I argue that the relationship between speech tone and Yorùbá music is far from incidental and aesthetic; text unambiguously determines the structure of Yorùbá music, particularly in traditional forms that have been resistant to European influence.

---

1 This paper has benefited greatly from my presentation of an earlier version at a three-day multi-disciplinary workshop at the Phonogrammarchiv & Institut für Corpuslinguistik und Texttechnologie titled, "Relationships of Speech Tone and Music," July 5-7 July, 2012. The conversations that took place with Murray Schellenberg, Robert Ladd and others, as well as suggestions from the two blind reviewers, have made a significant contribution to this publication.

2 See, for example, Schellenberg (2012) on Fanti (268), Ewe and Twi (269, 273) from Ghana, Xhosa (269) and Shona (271), and Blacking (1967: 166–190) on Venda in Southern Africa.

3 The term "tone" is ambiguous in a cross-disciplinary context across music and linguistics as in musicians' vocabulary it indicates: 1) sound production or timbre; or 2) a kind of interval (the distance between two notes). In the terminology of linguists, "tone" indicates relative pitch. To reduce this ambiguity, I will use: 1) "note" or "pitch" to indicate the fundamental pitch frequency of an instrumental or sung note; 2) "pitch" for the fundamental of an utterance in natural speech; 3) the U.S. American musical terminology "whole-step" (as opposed to British "tone") for the musical interval; and 4) "speech tone" to indicate the relative pitch of a spoken syllable.

In ethnomusicological literature about the boundaries of speech and song (e.g., Herzog 1934; List 1961, 1963; Bright 1963; Agawu 1995), various vocal forms have been theorized as a linear progression of speech → chant → song, whereby chant is regarded as "heightened speech" and song is farther from natural speech than chant. In this article, I argue that Yorùbá chant is closer to natural speech than song only in terms of its rhythmic organization, which is closer to free speech time than the regular, danceable rhythm of song. Chant generally employs a restricted pitch structure that distinguishes it from everyday speech. Furthermore, some forms of chant frequently "flatten out" the spoken tones, as in, for example Ọ̀yọ́-style Ṣàngó *pípè*.[4] Conversely, song is closer to natural speech than chant in respect of its pitch organization, which exaggerates the pitch contours of speech, some of which are so fast that they are frequently difficult to perceive. As space does not allow me to detail my findings about the pitch and temporal organization of Yorùbá chant in this article, I will briefly summarize the broad properties of different kinds of Yorùbá chant and focus primarily on how text is realized in the Yorùbá song style of the Ọ̀yọ́ people.

By examining instances where the pitch contours of vocal music move in unexpected ways or even appear to deviate from spoken speech tone, my guiding research questions have been: 1) why does melody apparently differ from speech tone in song melodies?; 2) what are the guiding principles underlying these "deviations?"; 3) where in the musical and/or linguistic construction is this apparent diversion of relative speech tone and melodic contour likely to occur?; and 4) how consistent are such tone/pitch diversions? As a foreign speaker of Yorùbá and a music analyst, my research has been led by hunches about how Yorùbá pitch contours behave in musical practice, as opposed to theory. My own observations about the shaping of melodic contours along with input from knowledgeable colleagues have directed me to readings in linguistics that have confirmed and explained my conjectures.[5]

The Yorùbá and their Language

The homeland of the Yorùbá people straddles contemporary southwest Nigeria and southern Benin. The ethno-linguistic term Yorùbá has been employed since the 1890s to encompass peoples who speak some twenty dialects across the region. With inadequate census numbers, there are said

4 Ìṣ̀ọ̀lá (1975: 789–790) describes how low tones (and very occasionally high tones) are frequently assimilated to mid tones in this style, resulting in chanted text dominated by mid tones.

5 In particular, I am grateful to 'Túndé Adégbọ́lá, who offered valuable input on an earlier draft of this paper and recommended key texts.

to be twenty to thirty million speakers of Yorùbá dialects. Different dialectal systems result in contrasting musical tonal systems (pers. comm., Túnjí Vidal, July 2011, and Ọlábọ̀dé Ọmọ́jọlà, November 2010). The current study describes the use of Ọ̀yọ́ dialect (regarded as Standard Yorùbá) using fieldwork data collected in Ọ̀yọ́, Òṣogbo and Ẹ̀rìn-Ọ̀ṣun since 1999.

Yorùbá is comprised of eighteen consonants, seven oral vowels, five nasal vowels plus a syllabic nasal. Most of the consonants correspond with English, with two exceptions: the voiceless labio-velar stop *p* /k͡p/ and the labio-velar stop *gb* /g͡b/. From what I can ascertain, only the consonant *r,* a tongue flap, influences the rhythmic organization of Yorùbá vocal music, as I will exemplify below. Yorùbá's vowel sounds are as follows: *a* /ɑ/, *e* /e/, *ẹ* /ɛ/, *i* /i/, *o* /o/, *ọ* /ɔ/, and *u* /u/. Additionally, Yorùbá has five nasalized vowels: *an* /ã/, *ẹn* /ɛ̃/, *in* /ĩ/, *ọn* /ɔ̃/ and *un* /ũ/, plus a syllabic nasal /n/. Nasalized vowels do not appear to have an influence in musical setting, whereas syllabic nasals increase the duration in the same manner as a vowel.[6]

Yorùbá is a true tone language with three relative pitch bands. Spoken Yorùbá frequently glides between vowels, as in the following examples: 1) within a single word (as in *dáa*); 2) from a word ending in a vowel followed by a word that begins with a vowel (as in *ṣé ìyá*); or 3) from a word ending in a vowel followed by a word that is comprised of just a vowel (as in *ṣó o*). The following example demonstrates how changes in relative tone can change meaning, and indeed may create opposite meanings:

A màá lọ 'we will definitely go' – *mà* indicates 'definitely';
A máa lọ 'we will go' – *máa* is a future prefix;
Má(à) lọ 'don't go' – *máà* denotes the imperative negative.[7]

Most Yorùbá verbs end with a vowel and most nouns start with a vowel, therefore verbs and nouns are frequently contracted into a single word. Thus in *kà ìwé* (*kà* 'read' + *ìwé* 'book', *kàwé* 'read' ('a book' is implied).[8]

6 As well as subdotting the vowels *ẹ* and *ọ*, the consonant *s* /s/ may also be subdotted to indicate /ʃ/. Yorùbá high tones (H) are marked with an accent acute, mid tones (M) have no marking, and an accent grave is used for low tones (L).

7 The low-tone *a* here is not pronounced in standard Yorùbá (Ọ̀yọ́ dialect), but the following word is pronounced as if preceded by a low tone. If the next word starts with a high tone, the syllable rises steeply, a mid tone is slightly flattened, while a low tone is unaffected (Barber & Oyètádé 1998: 40).

8 In such cases where a word ending in a vowel is followed by a word beginning with a vowel, one of the vowels may either be deleted or assimilated, depending on the construction. If a vowel is deleted, this affects the duration as the vowel, which takes time to utter, is simply not there anymore. If a vowel is assimilated, its duration may still be there and it will sound like the vowel to which it was assimilated (Barber & Oyètádé 1998: 20).

Additionally, prepositions are constantly elided with nouns. For example, *ní ilé* 'at home' becomes *nílé.* Consonants can also change when elisions are instigated, such as in *ní oògùn* becomes *lóògùn* 'have medicine', or 'able to cast spells.' There are complex rules about which vowel and tone band takes precedence when two vowels are combined.

Additional to the many kinds of elisions, there are numerous phonological aspects of Yorùbá that do not appear in the orthography (apart from in the pedagogical literature designed for foreigners). Consequently, the language frequently sounds significantly different to how it looks, particularly in terms of relative tone and duration. For example, when two verbs are juxtaposed, the first one is extended (but not so in the orthography). So, *mo fẹ́ lọ* is pronounced *mo fẹ́ẹ́ lọ* 'I want to go.'

Shifts in relative speech tone may also mark grammatical function. For example, nouns are extended with a tone glide when followed by an adjective or other modifier. Thus, *ilẹ̀* Yorùbá 'Yorùbáland' is pronounced *ilẹ̀ẹ̀* Yorùbá. Similarly, possessive pronouns are marked by a glide onto a mid or low tone: *ìwé mi* 'my book' is pronounced *ìwéè mi.* A (proper) noun that ends on a low tone adds an ascending glide before a high tone affirmative verb. Thus, *Bọ́sẹ̀ jí* is pronounced *Bọ́sẹ̀ẹ́ jí* 'Bọ́sẹ̀ woke'. After a low-tone verb, mid-tone object pronouns shift to a high tone. For example, *Mojí ń nà mi* 'Mojí is hitting me' is pronounced *Mojí ń nà mí.* Furthermore, a mid-tone pronoun before the continuous marker *ń* shifts to a low tone. For example, *mo ń lọ* 'I am going' is pronounced *mò ń lọ.* The following example makes explicit this difference between how Yorùbá is usually written and how it is pronounced. Pronounced extensions are marked with bold text, contractions are underlined in both versions, and unitalicized text denotes a change of consonant, vowel and/or tone:

Written:

Mo fẹ́ lọ sí ilé ṣùgbọ́n mi ò rí kọ́kọ́rọ́ mọ́tò mi. Mi ò lè gun kẹ̀kẹ́ mi nítorí pé ó bàjẹ́. Mo ti pe ìyá mi ṣùgbọ́n kò sí ní ilé.

Pronounced:

*Mo fẹ́**ẹ́** lọọlé ṣùgbọ́n* nǹ *rí kọ́kọ́rọ́**ọ*** *mọ́tò**o** mi. Mi ò lè gun kẹ̀kẹ́**è*** mí *n**í**torí pé ó bàjẹ́. Mo ti peyá**à*** *mi ṣùgbọ́n kò sí ńlé.*

'I want to go home but I can't see my car keys. I can't ride my bicycle because it is broken. I just called my mother, but she is not home.'

These examples of shifting tones and vowel glides – not reflected in the orthography – are nowhere near exhaustive. They are presented as examples in order to demonstrate that there are many instances where the contour of

a sung melody or of drummed speech may appear to be at odds with the tones when viewing transcribed music notation and text, but the musical pitch contour may actually follow, or frequently exaggerate the correctly pronounced utterance, not fully represented in writing.

Calling Yorùbá a "three-tone language" creates a conceptual illusion that there are only three absolute pitches. Naturally, linguists are not vulnerable to this illusion, but for non-linguists studying speech and music, the misconception of three fixed pitches appears to creep into the analyses, even in those of native speakers of Yorùbá. The use of a three-tiered *solfège* system of *do*, *re*, *mi* is used among the Yorùbá to indicate low, mid and high tones respectively. *Solfège* was first employed by British missionaries, who documented various forms of Yorùbá oral literature from the mid nineteenth century.[9] This system has been internalized by the Yorùbá and is enduring and is still employed for both language and musical learning in Nigerian classrooms and pedagogical textbooks. According to what I call "*do re mi* consciousness", Yorùbá language and song analysis frequently employs three-line staff notation (see Figure 1), which is not representative of the full tonal range of performances and limits a retrospective analysis of published sources.

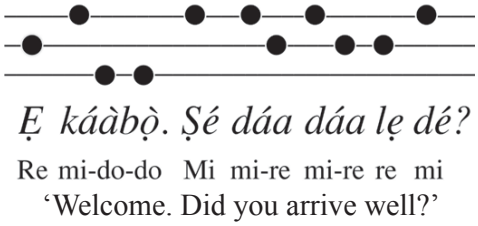


Figure 1: The use of *solfège* and a three-line staff for Yorùbá.

9 Many Church Missionary Society (C.M.S.) missionaries were themselves Yorùbá or Saros (Yorùbá slaves repatriated to Freetown, Sierra Leone, after abolition in 1807). The C.M.S. founded Fourah Bay College in Freetown in 1827, with the view of training Africans as teachers, schoolmasters, catechists and clergymen in order to promote education and the spread of Christianity in West Africa. Fourah Bay College was later affiliated with Durham University from 1876–1967. Yorùbá C.M.S. missionaries received British education there, which included formal (European) musical instruction (see Marcuzzi 2005: 359 and Peel 2000: 288–295). Many of these educated Saros found their way back to Nigeria and became part of an elite, educated class. Exhibiting an ethnocentric view as ethnic insider but cultural and certainly musical outsider, and also providing evidence that European musical pedagogy was employed, Yorùbá missionary James White commented in his journal, "The fine arts of the Yorùbá nation, of which music is a part, are in a very rude and unfinished state […] While meditating on what way sacred music can best be taught, so that men and women and adults and children can unite in heart and voice, the thought came into my mind, that the Ọtas are spoken of as superior to the [other] tribes of the Yorùbá nation in these things […] Our converts, when heathens, certainly had hymns and songs of praise in honour of their gods – might they not also, now that they are Christians, compose songs and hymns of the GOD of gods and LORD of Lords?" He goes on to comment that erroneous and unscriptual expressions were corrected, and that the Yorùbá paid no attention to rhythm or metre, and had not been able to make two verses of the same length or number of syllables (James White, journal, 31 December 1857, C.M.S. Archive, CA2 (87).

As I will demonstrate towards the end of this article, the five-line staff is also inadequate for a fine analysis of Yorùbá vocal music as there is a broad margin of pitch tolerance in Yorùbá song, which routinely employs microtonal shifts. However, for the purposes of this study, conventional staff notation provides easy access to the analysis.

Yorùbá Oral Literature

Tone carries a high lexical load in Yorùbá; the melodic contour of sentences allows the listener to extract much of the meaning. The success of Yorùbá drum language attests to the significance of relative pitch (see Euba 1990 and Villepastour 2010). Accordingly, song melody (*ohùn* in Yorùbá) is informed by the pitch contour of natural speech, though routinely seems to "deviate" from what one would expect when based on the theory of three tones and their representation in Yorùbá orthography. Unlike in Yorùbá drumming traditions such as *bàtá*, where the relative pitches of speech tones are played inflexibly on fixed-pitch drum heads, relative speech tone in song melody is not limited by fixed pitches and is mobile in much the same way as the pitch bands are mobile in natural speech.

The vocal repertoires I draw from in current research derives from the pre-Islamic and -Christian spiritual tradition of the Yorùbá known as *òrìṣà* religion or *òrìṣà* worship. The various repertoires of *òrìṣà* devotion provide an interesting data pool for text in music studies as they are less prone to (though not entirely free of) the colonial musical influences forcefully present in Yorùbá church repertoires and popular music. It is arguable that only Yorùbá people who are schooled in the missionary system and enculturated into *do re mi* consciousness, be they Christians, Muslims or traditionalists, ever sing text contrary to speech tone direction, or, in tune with a tempered scale. A deeply embedded and largely unconscious European influence has pushed song melodies towards tempered tuning, heptatonic scales (as opposed to pentatonic scales in most traditional music) and melodic forms obeying European rules of functional harmony. Indeed, one can sometimes hear a leading note (7th degree of heptatonic scales) creep into otherwise pentatonic forms in the *òrìṣà* song repertoire.

There are numerous forms of Yorùbá oral literature that are situated on the continuum between speaking (*ọ̀rọ̀ sísọ*) and singing (*orin kíkọ*). Some of the common forms that are usually classified as chant or heightened speech (for which there is no collective Yorùbá term) are *ìyẹ̀rẹ̀-Ifá* (chanted by Ifá divination priests, *babaláwo*), *ìwì* (poetry of masqueraders in the Yorùbá ancestral Egúngún cult), *ìjalá* (oral literature of hunters), *rárà* (the royal bards'

chant), and more specifically *rárà ìyàwó* (the bride's lament).[10] Denoting some of the differences in both function and vocal timbre of these forms, the verbs that precede each genre in place of "perform" are specific: *ké* 'cry, shout out' *ìwì*, *sun* 'secrete' *rárà*, *pè* 'call' *ìjalá*, and *ki* 'recite' Ifá. Aside from their discrete functions and expressive qualities, the text content of chants may be very similar across various chant genres. *Oríkì* is the generic form of chant that provides the text content and hence raw materials for all of the various styles of Yorùbá chant. Other stylistic features such as range (voice placement), tessitura, internal intervallic structure, relative phrase length, and the ratio of indeterminate pitch (spoken) to pitched (sung) utterances also determine the discrete classification of chants. Using a melograph, Vidal (1978) identified the typical pitch registers and melodic ranges of distinct chant forms, usually around a perfect fourth or fifth (see Figure 2).

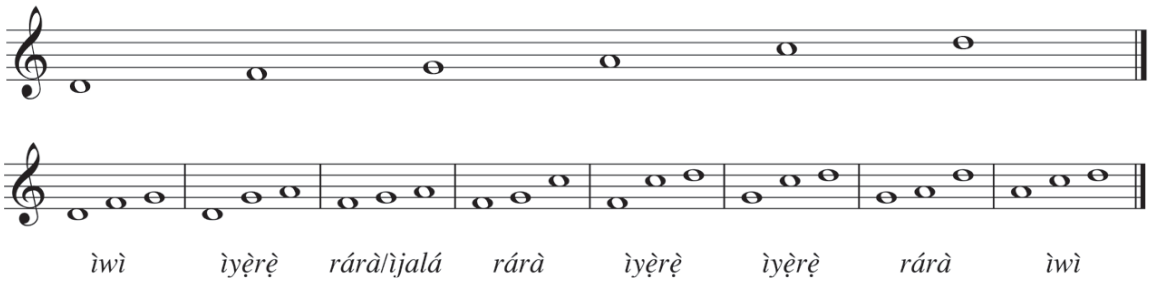


Figure 2: Typical registers and ranges of various Yorùbá chant forms.

The above notation gives examples generated from a pentatonic scale with a tone centre of D (as notated in the top staff), but the cells illustrated would not necessarily fall on these absolute pitches. Their internal intervals (which are highly restricted relative to song) are dominated by thirds, fourths and fifths, and there is an absence of a regular, danceable pulse (as is present in song).[11]

10 For scholarship on the various forms of Yorùbá oral literature, see for example: Gbàdàmọ̀sí & Beier (1959); Babálọlá (1966); Welch (1972, 2001); Ìṣọ̀lá (1973); Abímbọ́lá (1975, 1976, 1977); Barber (1991); and Bascom (1991).

11 In order to communicate the typical internal structures of the tonal system I have observed in Ọ̀yọ́ song styles, I employ European terminology such as "thirds" and "fourths" in the absence of any equivalent emic terminology.

The main differences between Yorùbá chant and song are as follows:

Chant	Song
Free speech time, exaggerated phrase structure	A regular pulse, more symmetrical rhythms
A constricted, idiomatic timbre	A more open, relaxed timbre
A restricted intervallic structure	A more fluid intervallic construction
A restricted tessitura, around a perfect 4th or 5th	A larger tessitura, around a major 9th

Table 1: Distinguishing features of Yorùbá chant and song.

Chant typically employs three-note cells for substantial stretches of a recitation, frequently shifting to other three-note cells, often after pauses. The three-note cell may shift back to the original one or to yet another cell. Figure 3 shows the pitch cell-syllable relationship in an excerpt of an *oríkì* performed by Ṣákìrat Àyándòkun iṇ Èrìn-Ọ̀ṣun, July 2011. The note head with the cross indicates spoken syllables. The last three syllables in this excerpt were low spoken notes, which frequently occur at cadence points.

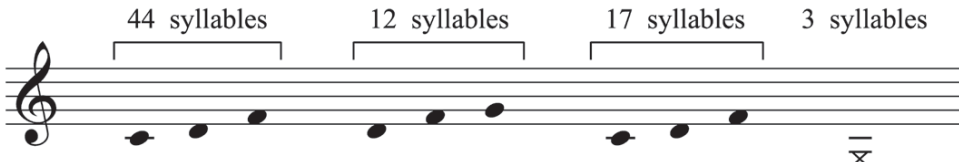


Figure 3: Pitch cells in an *oríkì*.

Vidal (1978) refers to such shifts as "modulations". To avoid inadvertently alluding to pre-determined shifts in tone centre, I use the term "cell shift", whereby the range and register shifts up or down to different three-note cells. This musical structure appears to be directly aligned to the phenomenon in natural speech that linguists call "terracing". Such cell shifts in songs tend to be much more frequent than in chant and cover a larger range, hence imitating the tone terracing of speech more closely than chant does.

Language in Yorùbá Song

After summarizing some recent cross-cultural research, Schellenberg (2012: 268) argues, "Changing a few tones in an utterance does not appear to make any difference to comprehension." This may be true for many true tone languages, but as African language polyglot and analyst Kubik (2010: 158) reminds us,

> Yorùbá is a language in which tone is used to extreme degrees, and therefore it has been quoted frequently as a classical example of the nature of West African tonal languages. […] Lack of awareness or the inability to use the tonal principles properly may catapult a non-Yorùbá visitor into comic situations.

The British missionaries' endeavour to superimpose their Christian musical repertoire onto the Yorùbá in Nigeria resulted in a systematic annihilation of their own hymns, often with a humorous outcome. By translating English hymn texts into Yorùbá and then methodically setting them to their original British melodies, the results would usually be unintelligible, but at times generated random Yorùbá texts. For example, the lyrics of the hymn "Oh Come All Ye Faithful" were translated into Yorùbá as *Wá ẹ̀yin olóòótọ́* 'Come all of you blind followers', but when sung to the original melody, the text became *Wa ẹ yìn o lóò tọ̀* 'Come and dig for palm kernels ye that are fond of passing water' (translation Adégbìtẹ́ 1978: 99):

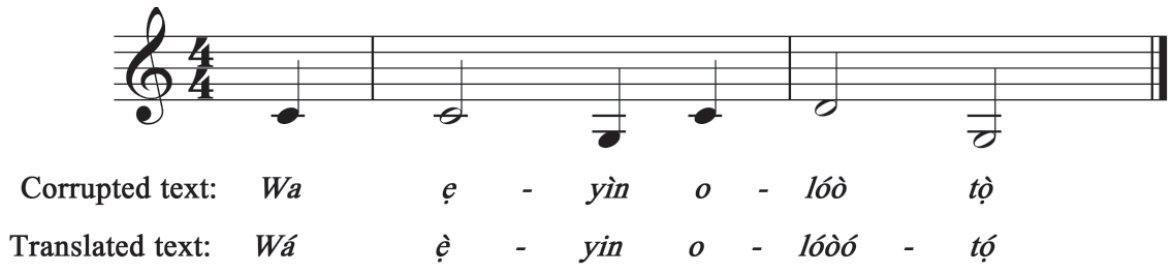


Figure 4: A British hymn translated into Yorùbá.

Some Yorùbá scholars, such as Akin Euba, assert that one can sing against the tonal structure and be understood,[12] but Euba himself is a product of missionary school education and was brought up in Lagos, where the musical terrain is dramatically different from the areas in which I have worked, where certainly elderly devotees are illiterate and have had little influence from Christian forms and popular music. In support of Yorùbá scholars such as Vidal (1978) and Adégbìtẹ́ (1978), who also insist that melodic contour must follow speech tone, I will demonstrate how Yorùbá speech and melody contours cohere in the more traditional repertoires.

Songs of the Ọ̀yọ́ Yorùbá generally employ two kinds of pentatonic scale (confirmed by Vidal 1978: 14) (see Figure 5), though neither conform to the European tempered scale but contain microtonal variations.

12 Euba stated this in a filmed interview with 'Túndé Adégbọ́lá, kindly shared with the author.

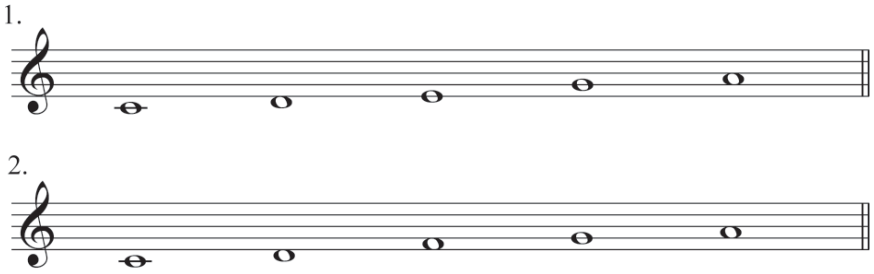


Figure 5: Pentatonic scales of the Ọ̀yọ́ Yorùbá.

Within either of the Ọ̀yọ́ pentatonic scales, the three-note (low-mid-high) speech tone cell can fall anywhere in the row:

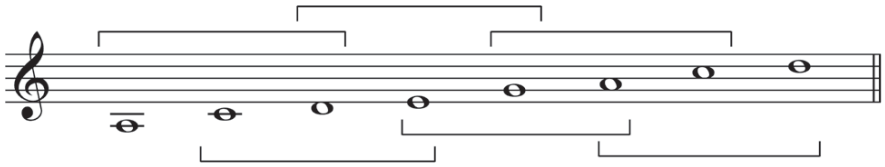


Figure 6: Melodic cells within the pentatonic scale.

Melodic cells also commonly miss a step in one or more places, often creating angular, large-interval melodies, depicted in Figure 7 (the crossed note head indicates a missed step):

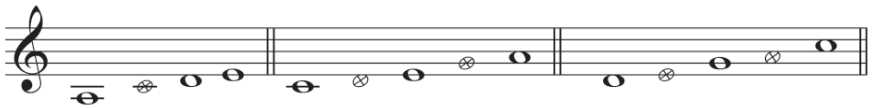


Figure 7: Melodic cells that miss a step in the pentatonic scale.

Due to these internal constructions, when Yorùbá language is sung, these shifting three-tiered cells create melodies within a pentatonic structure while broadly maintaining the relative pitch of the syllables in ordinary speech. These cells are usually asymmetrical rather than stepwise or triadic. In Yorùbá song, these three-tiered cells are in flux (as in speech). The pentatonic scale is an ideal receptacle for Yorùbá as the absence of half-steps disambiguates neighbouring speech tones. Indeed Vidal says that spoken Yorùbá would not normally place neighbouring speech tones less than a whole-step apart.

In his comparison of different chant forms, Vidal (1978) discusses the concept of "cadence", which he applies to the resting point of an utterance, as opposed to the musicological term used within functional harmony. Vidal uses the word to signify the "end portions of the poetic lines" (Vidal 1978: 15). He goes on to state, "the one that falls on the middle tone has the effect of a full resting point and is, therefore, referred to as the full cadence." He deems rest points on a high or low tone as "half cadences". Vidal's observations about Yorùbá chant significantly predate relevant linguistic research on the Yorùbá language (e.g., Akinlàbí & Liberman 2001) that explains how emphatic particles function. Emphatic particles are a kind

of clitic and come at the end of Yorùbá phrases and fall or rise at cadence points in natural speech. Yorùbá emphatic particles, which are sometimes referred to as "modifiers", are usually pronounced *o* (or can extend the "host" vowel that precedes it) at the end of a phrase or sentence. Emphatic particles are regarded as tonally neutral, so the convention is to write them without a diacritical tone mark, designating a mid tone in Yorùbá. But according to Akinlàbí & Liberman (2001: 41), "The [emphatic] clitic is realized as Low, except when the preceding tone is Low, in which case it is toneless (i.e. Mid)." Accordingly, in Yorùbá song, emphatic clitics can fall on mid or low tones at cadence points. If the utterance ends on a high tone, a mid or low tone clitic is usually inserted to ensure a dropping cadence.

The following song for the *òrìṣà* (Yorùbá deity) Ọ̀ṣun (recorded in Òṣogbo in 2002) demonstrates several of the principles discussed above.

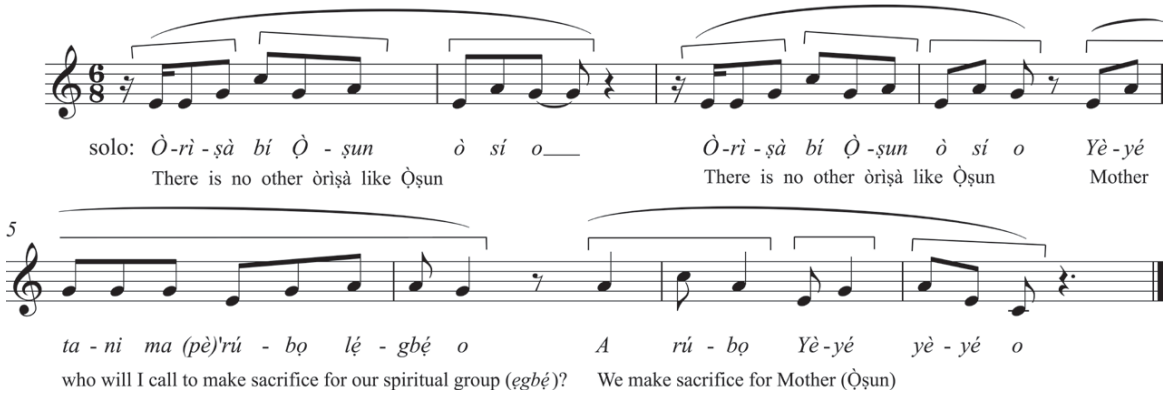


Figure 8: Song for the *òrìṣà* Ọ̀ṣun.

Rather than just employing three pitches of the three conceptual pitch bands (speech tones) as one might hear in chant, the relative speech tones are fluid, as in natural speech. Each shifting three-pitch cell is marked by square brackets above the notes. In respect of this kind of pitch arrangement, song resembles natural speech more than chant, the latter of which considerably restricts the choice of pitches and may employ just three or even two pitches for long stretches of text (as described by Ìṣọ̀lá 1975: 789–790). The shifting three-pitch cells observed in Figure 8 are similar to "tone terracing" heard in natural speech, which Connell & Ladd (1990: 4) describe as follows:

> […] the repeated application of downstep (or, in some cases, upstep) produces frequency bands with the form of a sequence of steps or terraces, with the result that, at the end of an utterance with several downsteps, the realisation of H tone will often be phonetically lower than the utterance-initial realisation of L.

Discussing the specifics of tone terracing in Yorùbá, Connell and Ladd reported the findings of experiments in which they discovered that high and

mid tones may be downstepped after a low-tone in sequences containing HLH, resulting in HL↓H. For example, in the phrase *Ṣé àp**ọ́n** ò wọ ṣòkòtò* 'Did the bachelor not wear any trousers?', the third syllable (in bold) will be lower than the first one. Connell and Ladd also identified that downstepping is not triggered by LH alone, but is more associated with HL.

What happens in melodic rendering may be slightly different from natural speech, but one certainly sees a kind of downstepping and terracing described by Connell and Ladd. Rather than terracing triggered just by HLH, in this musical example, the HLM of *bí Ọ̀ṣun* in bar 1 is downstepped to three different pitches for the LHM of *ò sí o* immediately following in bar 2. One sees a further upward cell shift on *a rúbọ* in bar 6, followed immediately by a shift down.

This example (Figure 8) brings several other principles to light. One can see four emphatic particles *o* at phrase cadence points (bars 2, 4, 6 and 8). At odds with Akinlàbí and Liberman's observations about natural speech, it is only the last emphatic particle that appears to drop to a low tone, though its relative drop from a high tone could feasibly be mid. Regardless of whether it is mid or low, the emphatic particle alleviates the phrase from ending on a high tone and creates a dropping cadence, as in natural speech. Although the last three notes of the song appear to defy the speech tone, the song melody is likely exaggerating natural speech by falling significantly in pitch at the end of the utterance. Here, there is no ambiguity on the word *yèyé* at the end of the Ọ̀ṣun song, because it is a repeated word. Of this tendency to drop in pitch, Vidal (1978: 19) states, "This descending movement seems to reflect the speech tone contours of Yorùbá language, whereby the speaker in ordinary daily speech has a tendency to drop the voice at periodic points." Linguists refer to these dropping tendencies at the end of an utterance as "final lowering",[13] which may explain the phenomena Vidal wrote of, and that I see at work in chanted and sung texts. In terms of cadence points, Ladd and Connell corroborate Vidal's findings. They did a series of experiments in Yorùbá that measured these final lowerings, and established that they routinely occur in natural speech without violating intelligibility. They also discovered that the drops are steeper on repeated low tones, which I have found also to be true for song melodies. With Connell and Ladd's findings in view, the final melodic rendering of *yèyé o* is an exaggeration of natural speech, not an instance where the melody has accommodated the aesthetic needs of music, as other scholars may assert.

13 Connell & Ladd (1990: 2) explain final lowering as an "abrupt lowering confined to phrase and utterance ends".

Rather, it may be the aesthetic needs of speech that determine the fall in pitch, thus *adding* to intelligibility by punctuating the end of the utterance.

Some other interesting principles in melody setting arise in this short example. In the opening word, rather than three repeated low notes on *òrìṣà* as one might expect, we see a jump of a minor third en route to the high speech tone *bí* on the note C. This can be explained by the fact that the proper noun here adds a high tone before the affirmative verb so is actually pronounced *Òrìṣàá bí* but would not normally appear in the orthography. The choice of notes in bars 1 and 3 exaggerates this quick glide in natural speech. Additionally, one often sees musical variation on a repeated speech tone (especially where there are three in a row) to add musical interest, but this would not normally be more than a whole-step.

In bar 5, the sung text is redacted to *tani ma rúbọ́ lẹ́gbẹ́ o* but would be spoken as *tani ma rúbọ́ pè lẹ́gbẹ́ o* (*pè* means 'call'). I was initially puzzled by this anomaly until the performer, Doyin 'Fáníyì, added the word *pè* in her written text and English translation. The low melody note on *rú* stands in for this missing syllable, which may have been eliminated so the phrase could fit into the regular six-pulse rhythmic structure of the song. I have found such instances throughout the song, chant and surrogate speech drumming repertoire, where a missing or elided syllable is not uttered, but is represented by the (apparently deviant) relative pitch only. When taking these explanations of apparent tone deviation in this section of an Ọ̀ṣun song into account, there are, in fact, no deviations from the broad relationships of speech tones one would find in spoken Yorùbá.

One more point that comes up in this example is the musical rendering of the consonant *r*. In drummed speech surrogacy, the syllable before the tongue flap *r* is temporally contracted, and we see that here in the rhythmic setting of the word *òrìṣà*, which is always performed with a shorter first syllable in drum language.

The song shown in Figure 9 for the *òrìṣà* Ọbàtálá reinforces some of the above points. I have marked out where the melodic cells shift with brackets over the staff. One can see a series of cadence points marked with arrows in bars 2, 4, 7, 8 and 10).

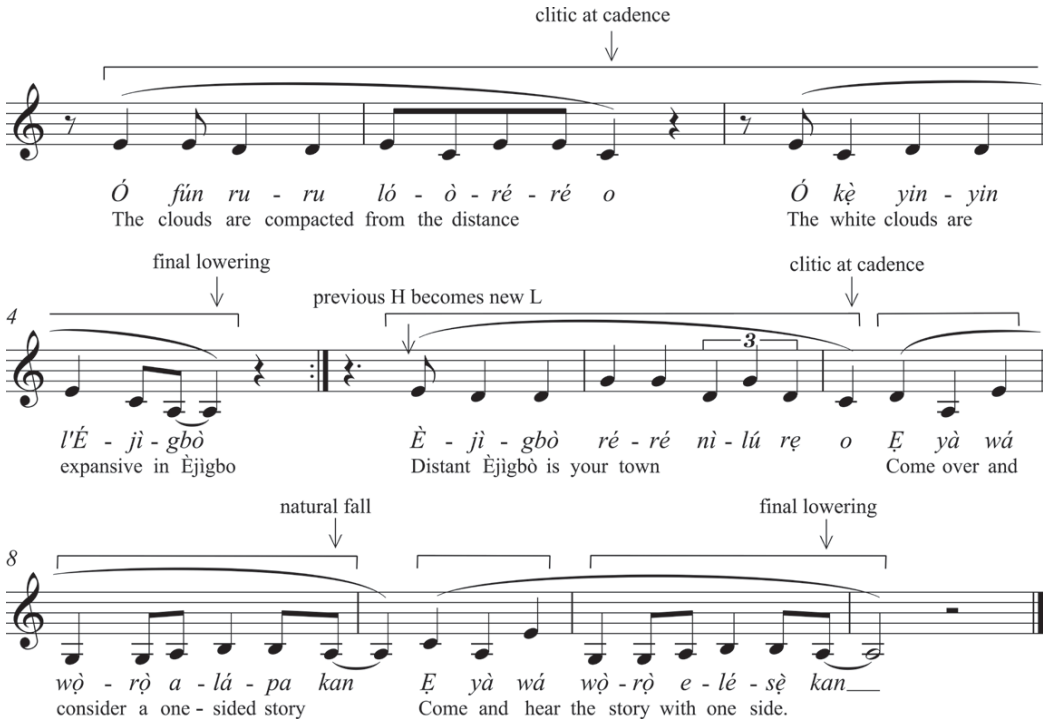


Figure 9: Song for the *òrìṣà* Ọbàtálá.

Only one of these cadence notes (bar 8) falls where one would expect based on the written speech tones and only one (bar 10) apparently deviates from the speech tone direction at the final cadence. The latter instance corresponds with the fall in pitch at the end of a Yorùbá utterance (final lowering), regardless of speech tone, at what Vidal calls the "full cadence". On the last note in bar 2, according to Akinlàbí and Liberman's theory, the emphatic clitic *o* should actually fall on a low tone, while the last low tone syllable *gbò* has an exaggerated drop in pitch at the next cadence point in bar 4, emulating final lowering in natural speech. In bar 5, the note E was previously a high tone (as in bar 1 on *Ó fún ruru*, bar 2 *lóòréré*, bar 3 *ó*, and bar 4 on *l'Éjìgbò*) but now becomes the new low tone at the beginning of bar 5 as the pitch generally lifts (or terraces upward) here, perhaps initiating a higher cell to add emotional intensity. Immediately following in bar 5, we see the shift of a whole-step down to D on the second and third syllables of *Èjìgbò*, emulating the downtrend[14] one would find on repeated low tones in natural speech while adding musical interest. The downtrend on both renderings of *Èjìgbò* in bars 4 and 5 fits with Connell and Ladd's Yorùbá experiments published in 1990, which established that the cadential fall on repeated Yorùbá low tones is steeper than the downtrend on other tones. The downward contour of the melody at the cadence point in bar 4

14 Connell & Ladd (1990: 2) explain downtrend (also known as declination) as "any overall lowering of speaking pitch during the course of an utterance".

is significantly steeper than that of bar 5. We see another drop to a low tone emphatic particle (marked conventionally as a mid tone) in the first note of bar 7, and on the last note in the final cadence in bar 10, we see the only instance where the melody note apparently goes in the opposite direction to the written speech tone, thus emphasizing the final statement, as one would hear in natural Yorùbá speech.

Further to the above explanations, there are several discrepancies between the orthography of Figure 9 and what would actually be pronounced. *Nìlú rẹ o* (bar 6) exemplifies a phenomenon that is described in Yorùbá linguistics as the assimilated low tone. The actual sentence without elisions is *ni ìlúù rẹ̀ẹ̀ o,* but what is actually spoken is *nìlúù rẹ̀ẹ̀ o* with an extra low tone *ẹ̀*. This probably explains why *rẹ* is on a D (the same as the previous low tone), as the melody note stands in for both of the low-tone glides in *nìlúù* and *rẹ̀ẹ̀.* Furthermore, *alápa kan* is pronounced *alápaá ọ̀kan,* as realized by the melody. The second two syllables of *alápaá* use the same (high-tone) note B, while the low note A stands in for the elided *ọ̀* in *ọ̀kan.*

Additionally, Connell and Ladd found that "the sequence HLH, whether sentence initial or preceded by L, triggers downstep". We may see something similar (though not identical) at work in this song for Ọbàtálá in bars 2, 6 and 9-10, although a larger body of data is needed to substantiate this.

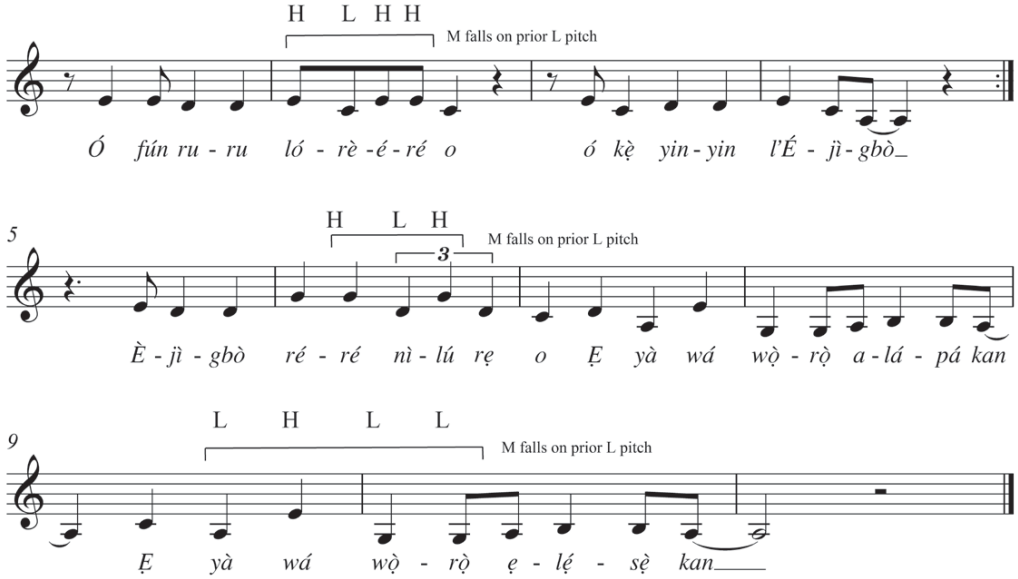


Figure 10: Tone terracing in song for Ọbàtálá.

The last note in bar two after the HLH sequence may be a low tone emphatic particle marked as a mid tone. While the emphatic clitic in bar 2 apparently falls on a low tone, in bar 6, one expects *rẹ* to fall on E between its preceding low and high tones, but it falls on a D, previously a low tone note, towards the cadence point on the low tone emphatic particle *o.* Another possibility is that *nìlú rẹ* is actually pronounced *nìlúù rẹ,* so the lowering of the melody

note reflects the unwritten low tone note after the HLH sequence in bar 6. The phrase *Ẹ yà wá wọ̀rọ̀ alápaá kan* is downstepped as in natural speech, whereby the first mid tone syllable *Ẹ* in bar 7 is lower than *È*, which started the previous phrase in bar 5. What is more explicit across all of these examples is that a phrase never ends on a high speech tone, or high musical note. Where the last word does end on a high tone, an emphatic particle is added to create a cadence point on the mid or low tone, as would happen in natural speech, or the text is over-ridden by final lowering, as in the final cadence.

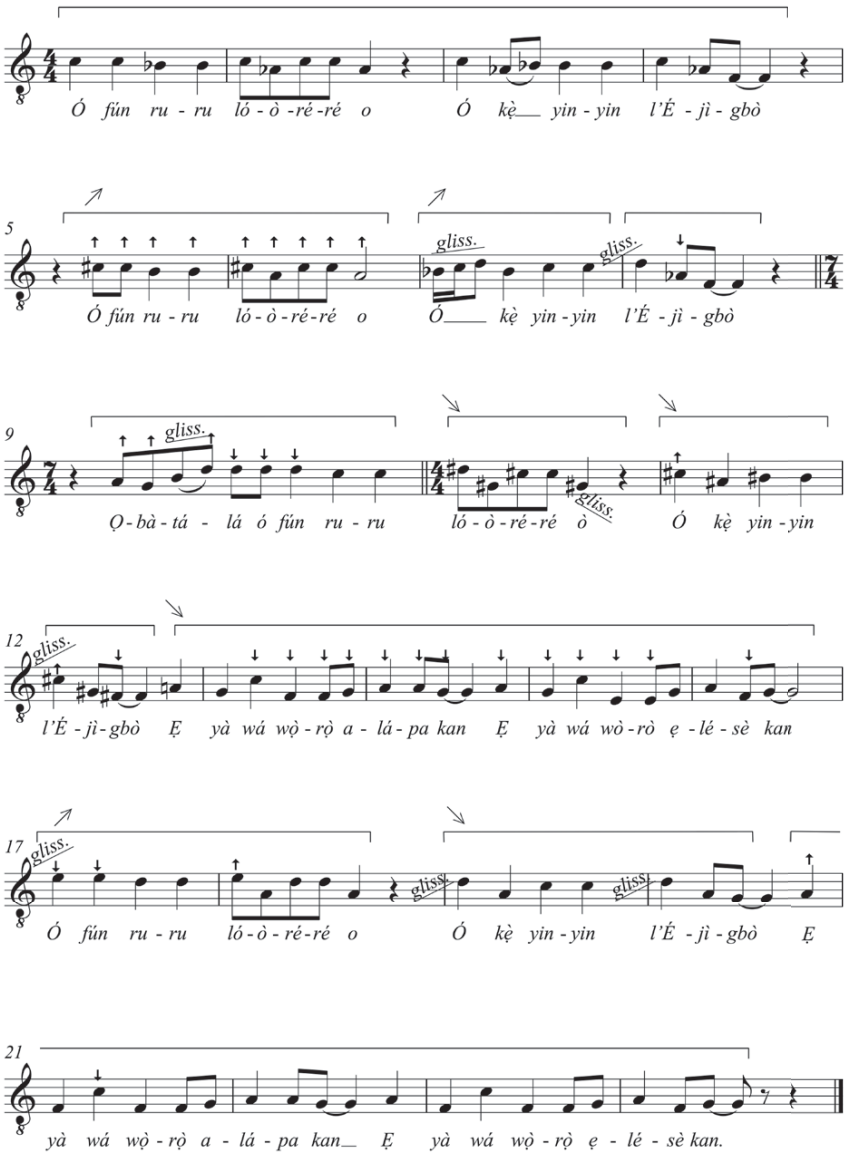


Figure 11: Transcription of Ọbàtálá song showing tonal shifts.

Beyond the points made above, the transcribed version of this Ọbàtálá song I have been using for my analysis has been modified to make the comparison of speech and melody more accessible. By "modified" I mean that the pitch was "corrected" to the nearest note from the tempered scale, as represented by staff notation. Likewise, contractions and expansions in

the regular pulse were ignored in order to work the transcription into a metric framework, so as not to distract from analysis of the speech tone and melodic contour. "Tidying" the transcription in the manner described here does not interfere with the pitch contour as such, but it certainly sidelines some larger musical issues. Furthermore, it obscures my argument that song is a form of heightened, exaggerated speech.

A more representative transcription is presented in Figure 11 from a different performance of the same song. This transcription reflects what one would be more likely to hear in a ritual setting. The arrows above the staff indicate tonal shifts, some of which are microtonal (less than a half step) while others are up to a whole step up or down, demonstrating the wide margin of pitch tolerance in traditional Yorùbá song. The changes in time signature indicate the metric flexibility one can expect in unaccompanied song in the same genre. This frequency and metric "untidiness" supports the notion that song is a form of heightened speech among the Yorùbá, rather than a neat musical format that violates the rules of speech tone and periodicity in natural speech for aesthetic ends.

Conclusion

Despite the common assumption that Yorùbá song melodies sometimes deviate from the speech tones of natural speech in order to serve the aesthetic needs of music, my data suggest that song melodies in fact do not routinely deviate. In cases where the melody does appear to deviate, it is almost always part of a formula that highlights tendencies in natural speech, such as descending patterns at cadence points, missing syllables hidden in elisions, and implicit grammatical syllable extensions and/or tone glides which are not represented in standard orthography and may be hidden even to native speakers who have been schooled into *do re mi* consciousness. Reversing Schellenberg's (2012: 276) assertion that "Music is willing to accommodate language as long as this does not interfere with the music", I propose that Yorùbá music not only accommodates language, but makes some of the phonological features of spoken Yorùbá more explicit, thus heightening meaning rather than compromising it. In this respect, Yorùbá speech largely determines Yorùbá melody.

My research is in some respects rudimentary and has proceeded from musical analysis rather than detailed speech analysis. What is needed to take this research forward is: 1) substantially more musical data to test my theories; and 2) collaboration with phoneticians and phonologists to further explore how the behaviours of Ọ̀yọ́ Yorùbá correspond with musical utterances. Until more is known about the acoustic phenomena of spoken Yorùbá, what can be deduced from Yorùbá musical expressions remains limited.

References

Abímbọ́lá, ’Wándé (ed.). 1975. *Yorùbá Oral Tradition*. (Ife African languages and literatures series, 1). Ilé-Ifẹ̀: Department of African Languages and Literatures, University of Ifẹ̀.

Abímbọ́lá, ’Wándé. 1976. *Ifá: An Exposition of Ifá Literary Corpus.* Ibadan: Oxford University Press Nigeria.

Abímbọ́lá, ’Wándé. 1977. *Ifá Divination Poetry.* New York etc.: NOK Publishers.

Adégbìtẹ́, Adémọ́lá. 1978. *Oríkì: A Study in Yorùbá Musical and Social Perception.* PhD dissertation, University of Pittsburgh.

Agawu, Kofi. 1995. *African Rhythm: A Northern Ewe Perspective.* Cambridge: Cambridge University Press.

Akinlàbí, A. & M. Liberman. 2001. “The Tonal Phonology of Yorùbá Clitics”. In: Gerlach, B. & and J. Grijzenhout (eds.). *Phonology, Morphology and Syntax.* Amsterdam: Benjamins, 31–62.

Babálọlá, S.A. 1966. *The Content and Form of Yorùbá Ìjalá.* Oxford: Clarendon Press.

Barber, Karin. 1991. *I Could Speak Until Tomorrow: Oríkì, Women and the Past in a Yorùbá Town.* Edinburgh: Edinburgh University Press.

Barber, Karin & Akin Oyètádé. 1998. *Yorùbá Wuyì. Ìwé Kìíní (Book one): A Beginners’ Course in Yorùbá.* London: Hakuna Matata Press.

Bascom, William. 1991 [1969]. *Ifá Divination: Communication Between Gods and Men in West Africa.* Bloomington & Indianapolis: Indiana University Press.

Blacking, John. 1967. *Venda Children’s Songs: A Study in Ethnomusicological Analysis.* Chicago & London: The University of Chicago Press.

Bright, William. 1963. “Language and Music: Areas for Cooperation”. *Ethnomusicology* 7: 26–32.

Connell, Bruce & D. Robert Ladd. 1990. “Aspects of Pitch Realisation in Yorùbá”. *Phonology Yearbook* 7: 1–29.

Euba, Akin. 1990. *Yorùbá Drumming: The Dùndún Tradition.* (Bayreuth African Studies Series, 21/22). Bayreuth: Breitinger.

Gbàdàmọ̀sí, Bakare & Ulli Beier. 1959. “Poetry of Masquerades: Four Poems of Egúngún Masquerades (*Ìwì*)”. *Odu* 7: 38–40.

Herzog, George. 1934. “Speech Melody and Primitive Music”. *Musical Quarterly* 104: 452–466.

Ìṣòlá, Akínwùmí. 1973. *Ṣàngó Pípè: One Type of Yorùbá Oral Poetry.* MA thesis, University of Lagos.

Ìṣòlá, Akínwùmí. 1975. “The Rhythm in Ṣàngó Pípè”. In: Abímbọ́lá, ’Wándé (ed.). *Yorùbá Oral Tradition.* (Ife African languages and literatures series, 1). Ilé-Ifẹ̀: Department of African Languages and Literatures, University of Ifẹ̀, 762–793.

Kubik, Gerhard. 2010. *Theory of African Music.* Vol. II. Chicago etc.: University of Chicago Press.

List, George. 1961. “Speech Melody and Song Melody in Central Thailand”. *Ethnomusicology* 5: 16–32.

List, George. 1963. “The Boundaries of Speech and Song”. *Ethnomusicology* 7: 1–16.

Marcuzzi, Michael D. 2005. *A Historical Study of the Ascendant Role of Bàtá Drumming in Cuban Òrìṣà Worship.* PhD thesis, York University (Toronto).

Peel, J.D.Y. 2000. *Religious Encounter and the Making of the Yorùbá.* Bloomington & Indianapolis: Indiana University Press.

Schellenberg, Murray. 2012. “Does Language Determine Music in Tone Languages?” *Ethnomusicology* 56 (2): 266–278.

Vidal, Ọlátúnjí. 1978. “The Tonemic and Melodic Character of Yorùbá Principal Chants”. Paper presented in the Seminar on the Inter-Relationship of the Arts in Nigeria, University of Ifẹ̀.

Villepastour, Amanda. 2010. *Ancient Text Messages of the Yorùbá Bàtá Drum: Cracking the Code.* Farnham: Ashgate.

Welch, David B. 1972. *Aspects of Vocal Performance in Sango Praise-poetry and Song.* PhD thesis, Northwestern University (Evanston, Ill.).

Welch, David B. 2001. *Voice of Thunder, Eyes of Fire: In Search of Shango in the African Diaspora*. Pittsburgh, Pa.: Dorrance.

How Language Makes the Drumming Beautiful: The Case of the Dagomba Lunga

(David Locke)

Introduction

> One day I asked Alhaji Abubakari Lunna, my teacher of the dance-drumming heritage of the Dagomba people of Northern Ghana, "None of your students of drumming know Dagbani. The audience has no clue about Dagomba chieftaincy. Here in the USA, we regard your drumming as music. Under these circumstances, why do you insist we learn the Dagbani meaning of drumming?"
>
> Alhaji Abubakari replied, "It is the language that makes the drumming beautiful. The very thing you love about this tradition – the sound of the drumming – is a product of language. The music just sounds better when it is shaped by language."

Alhaji Abubakari taught me Dagomba dance-drumming nearly thirty-five years, from when we first met in Ghana in 1975 until his untimely death in 2007 at the age of 72. I studied the music and instrumental technique of the *lunga*, a pressure drum, and the *gung-gong*, a bass drum with snare.[1] Our collaboration resulted in a printed monograph, *Drum Damba* (Locke 1990), and an online monograph, *Dagomba Dance Drumming* (Locke 2010). Although Dagomba drummers use a pedagogical system among themselves, their system is not designed for teaching Africans beyond their drumming families, much less non-African cultural neophytes. Never having worked intensely with a foreigner before me, Alhaji Abubakari had to figure out how to teach me and I had to learn how to learn. Rather quickly, we achieved clarity on the fundamentals of *lunga* technique, i.e., its three main melodic pitches made with an open bounce stroke. But then followed several decades of frustration: Alhaji would say, "Not bad, pretty good, but I cannot hear the language in your drumming," and I would think, "But I am imitating what you are playing as closely as I possibly can."

Finally, in 1999, I formally studied Dagbani, the language spoken by the Dagomba people.[2] Although never approaching conversational fluency,

---

1 The player of the lead *lunga* calls the ensemble to begin, signals endings, and plays the drum talks that are associated with each item of repertoire. The answer *lunga* drummers respond by playing the phrase that is appropriate for the item of repertoire. The *gung-gong* drum provides the groove that animates dancers. Players wear the *lunga* drum over their left shoulder, holding it snuggly up into the armpit with a long scarf. Held in the right hand, a curved wooden stick with a flattened knob on the end is the drumbeater.

2 I studied with M. Dawuda Sulley, Language Center, University of Ghana. Language learning materials for Dagbani are minimal. We produced an edited version of Mr Sulley's self-produced workbook, complete with

I learned enough about the sound and structure of the language to perceive features in Alhaji Abubakari's drumming that had previously not come to my attention. I had a musical "Eureka moment". Pressure techniques that I had previously regarded as virtuosic embellishment I now heard as features essential to rendering language on the drum. Enthused, I rededicated myself to obtaining proficiency in the subtle art of increasing and decreasing the tension on the cords that connect the drum's two heads. Alhaji noticed the difference in my playing, "Now I can hear the talk!" In this paper I will share some of these important techniques for rendering spoken Dagbani on the *lunga* drum and point out how language has impact on the aesthetic beauty of the drum's music (see Fiagbedzi on the concept of "the aesthetic" in African music).

Brief Literature Discussion

Akin Euba's comprehensive study of Yorùbá *dùndún* drumming provides helpful structure for this topic (Euba 1990). The basic point made by Euba (1990: 298) about the influence of text on musical style holds for the Dagomba case:

> To the extent that the íyáàlu drum uses a text, its tonal style is based on the rules of intonation which Yorùbá speech customarily follows. [...] For the íyáàlu drummer, words constitute the raw material of musical structures, material which although rich in musical possibilities remains a mere function of speech communication until its musical potential is developed.

The *iyáàlú* drum stylistically reproduces two parameters of human speech, rhythm and intonation (Euba 1990: 194), but the Dagomba *lunga* achieves more realism by representing subtle phonological features such as velar fricatives, final nasal endings and tonal glides. I suggest that, due to three factors, the Dagomba *lunga* allows a closer resemblance between language and drumming than the Yorùbá íyáàlu: (1) the way the drum head is fixed to the drum's wooden body, (2) the way the drum's pressure cords are manipulated, and (3) the way the drum is struck.

Amanda Villepastour (2010) observes a contrast between the staccato sounds of Yorùbá *bàtá* drums, which fade quickly to silence, and the legato sound of *dùndún* drumming, which enables an unbroken link between successive notes that permits imitation of speech features like *portamento*, i.e., smoothly connecting notes of different pitch, and *glissando*, i.e.,

---

audio recordings; contact the author for copies. My knowledge of the Dagbani language remains regrettably weak. I am grateful to an anonymous review for insight into its phonetics, as well as valuable secondary sources (Bendor-Samuels, Hyman, Fusheini, and Orlawsky).

precisely sliding between specific pitches (Villepastour 2010: 17). The *bàtá* tradition uses *enà* (coded language) as an intermediate stage between speech and drumming. *Enà* consists of drum vocables that not only denote types of drumming sounds but make up a vocabulary of a secret language that is intelligible only to initiates of the òrìshà cults for whom *bàtá* drums are played (Villepastour 2010: 91ff.). The drummer does not try to play speech *per se* but rather plays its code-talking derivative.
Summarizing systems of speech surrogacy, Villepastour (2010: 90) contrasts the process of encoding with that of mimicry:

> If one is to imagine a surrogate speech continuum with encoding on one end (where one would place, say, Morse code) and mimicking on the other (where one would situate synthesized speech), the traditional bàtá would be situated toward the encoding side [...] and the dùndún would be situated towards the mimicking side.

I would place *lunga* beyond *dùndún* in its mimetic capacity.

Like Villepastour, Kofi Agawu has studied the relationship of music to West African languages (Ewe and Siwu) whose "resultant intonational contour" carries semantic valence (Agawu 1995: 34). Appropriate for his interest in rhythm, he calls attention to three qualities of language: "stress, quantity, and resultant patterns," i.e., loudness, duration, and stock figures. When discussing "stress languages" versus "tone languages" he notes that changes in tone need not imply a change in rhythmic accent. As in Yorùbá and Dagbani, "negotiating the gaps between tone levels by means of glides" is crucial in speech and speech surrogacy:

> Ewe has both falling and rising glides. Some are structural in the sense that they are built into the meaning of the word, while others are non-structural in the sense that they are brought on by the mood or circumstance of the speaker. (Agawu 1995: 38)

Enriching prior music-language studies that only follow the impact of speech tone on musical pitch, Agawu emphasizes the phonological and semantic significance of temporality, "speech tone and speech rhythm are fundamental, character-defining traits of the two languages" (Agawu 1995: 42). Moving along Agawu's proposed continuum of rhythmic expression from gesture to dance, the latent musicality of language becomes increasingly manifest (see Agawu 1995: 28, fig.1.1):

> Language is turned into song when its latent musical elements are made patent, when the implicit becomes explicit, when the hidden becomes manifest. (Agawu 1995: 60)

This paper draws attention to how the *lunga* drum articulates the musicality of phonological features of language. Without these details, I argue, the *lunga* not only fails to fulfill its function as a "talking drum", but its music lacks subtle nuances that are crucial to its aesthetic beauty.

Drum Construction

Drummers control a *lunga's* pitches by squeezing and releasing the cords that run between its two drumheads. Although the exact pitches of each drum depend on the dimensions of its wooden body and the thickness of its two goatskin heads, Alhaji Abubakari consistently intoned three fundamental pitch classes, separated by predictable melodic intervals: low-tone to mid-tone is a minor third, mid-tone to high-tone is a major second, and low-tone to high-tone is a perfect fourth.[3]

Whenever he would drum, Alhaji Abubakari invariably began by playing an invocation to the Creator for providing human beings with the natural world and imbuing us with sense to make good use of it (see Fig. 3 and Locke 2010, *Nawuni Mali Kpam Pam*). Drummers themselves transform carved wooden drum shells into *lunga* drums using goatskin, antelope skin, cane, and raffia. The hourglass-shaped body of a *lunga* drum has the geometric shape of two gently curving bowls connected by a hollow cylinder. A drum rim made of coiled-up cane wrapped in raffia grass fits over the circular opening on each end of the drum's body. Goatskin is sewn onto the rim to make each drumhead. Cords made of antelope skin are strung back and forth between the two heads, holding them in place and enabling the drummer to apply the pressure that controls the drum's pitches. Alhaji Abubakari astutely pointed out that the pressure ropes actually pull on the cane rim, whose subtle movement affects the tension of the goatskin. The cane rim is the crucial technology that enables the drummer's pressure technique to make the drum mimic the phonological nuances of the spoken Dagbani language.

Pressure Technique

Performance training suggests that the Dagomba drumming technique on *lunga* achieves a closer representation of speech than heard from Yorùbá *dùndún* drums (for a description of *dùndún* technique see Euba 1990: 147–149). The basic principle is the same in both musical traditions: an increase or decrease in pressure raises or lowers the pitch. The drum's melody comes from the player's ability to consistently intone three pitch classes; not only

3 Notated in staff notation as B-D-E, these pitches and intervals are easy to sing but difficult to play with consistent intonation on an actual *lunga* drum.

must the melodic intervals be constant but also the drummer sometimes must glide upward and downward between pitch classes. Furthermore, for the *lunga* drum to replicate the sound of spoken language, some notes must be approached from above or below. My breakthrough came when I finally realized that the pitch bends of the *lunga* correlated closely with the sound of spoken Dagbani.

As much as it is difficult to achieve precise intonation on a fretless stringed instrument, the playing technique on the *lunga* drum is challenging because consistent intonation requires precise and subtle control of the tension on the antelope leather ropes that run between the two drumheads. The thumb hooks through two or three of the double-course *lunga* ropes to steady the drum and provide a focal point for the wrist, whose inner surface touches the *lunga* so that the lower arm segment from the elbow to the wrist can apply pressure to the lower front section of the *lunga* ropes. The upper arm passively applies pressure to the back section of *lunga* ropes simply by virtue of the way the drum snuggly fits into the armpit. Motion in the elbow joint raises and lowers the forearm; motion in the shoulder joint moves the arm in and out. The arm gets the pressure roughly correct, but it is the wrist that makes the subtle adjustments for precise intonation: hyperextension increases pressure by pressing inward and upward against the *lunga* ropes; flexion moves the wrist away from the drum, thus decreasing pressure and lowering pitch. Drummers who play *lunga* frequently often develop a callus on the inside of their left wrist, an indication of the essential function the wrist has in playing technique.

Sticking Technique

The curved drumstick lies in the fist so that the handle rests against the knuckles and the shank is roughly perpendicular to the forearm. The first finger of the right hand encircles the handle and the extended thumb clamps the stick in place. The other three fingers of the right hand lie gently on the stick handle, helping to flick and guide the stick toward the center of the drum skin. At the moment of impact, the hand is flexible so that the stick can rebound off the drum skin. The stick head moves in an arching motion with the speed and amplitude of its path in space affecting the force it exerts on the drum skin. Hitting the center of the skin while holding the stick with a relaxed hand and arm is the player's goal.

In addition to "straight strokes," as Alhaji Abubakari would say, two other types of drum strokes help the drum emulate the phonetics of spoken language. To mimic several sounds in rapid succession, Alhaji Abubakari uses a two-stroke press-roll. When language mimesis calls for a contrast in loudness – usually a relatively quiet first stroke followed by a louder

second stroke – Alhaji Abubakari tends to turn his wrist toward his body so that the first stroke is played with the upper side of the stick head and then turn the wrist away from his body back to the normal position so that the second stroke gets full power. Although the drum language can be played without rolling or turning the stick, it is misleading to think of these strokes as "ornaments". Rather, these strokes enable the drummer to imitate the spoken language with more realism.

Vocables

Vocables, or "phonoaesthetic oral notation" (Euba 1990: 214), are a way to play the drum with your mouth. In the Dagomba drumming tradition vocables themselves have no lexical meaning in the Dagbani language, but to call them nonsense syllables would be misleading since they precisely symbolize the drumming music (see Table 1). Vocables accurately represent the pitch, timbre, melody, rhythm, dynamics, and articulation of notes in a drum phrase. When teaching a new drum phrase to American students, Alhaji Abubakari usually began with vocables. After we learned to chant the drumming phrase, he instructed, "Now make the exact same sound on the drum." In my experience, the vocables taught by Alhaji are rather consistently used by most Dagomba drummers, especially in their vowels.

For the *lunga* drum, the three pitch classes are distinguished by vowel sound: "a" for low pitch, "e" for mid pitch, and "i" for high pitch.[4] The way the drumstick hits the skin is indicated by the consonant at the beginning of the vocable: in most musical settings the vocable for bounce strokes starts with "d". Thus, the most widespread vocables for the music of the *lunga* drum are "da" for low-pitch stroke, "de" for mid-pitch stroke, and "di" for high-pitch stroke. In special musical situations, such as a relatively soft-sounding initial note in a figure, Alhaji Abubakari uses "z". When Alhaji Abubakari verbalizes a two-stroke drum role, which he usually plays with turned-stick technique, he initiates the syllable with the sound "gr-" or "kr-".[5] Vocables also represent the duration of a note and the way it ends. Notes of longer than normal duration end with the nasal consonant "n". The "n" sound also occurs on the final note in a musical unit (figure, motive, or phrase).

4 The author's knowledge of the phonetics of Dagbani in particular, and the science of phonetics in general, is admittedly modest. An anonymous reader points out that Dagbani is believed to have two semantic tones (high and low), not three (high, mid, and low). Nevertheless, as I understand it, the vocable system for *lunga* music shows three pitch classes.

5 An anonymous reader with greater knowledge of Dagbani phonetics than the author notes the absence of consonant clusters in Dagbani, which means that the schwa vowel is heard between the velar stop and the "r".

Vocables also indicate notes produced by pressure on the *lunga* ropes, i.e., those that represent the all-important glides between the different tone levels of spoken Dagbani. The sound "h" indicates notes produced by increased pressure. Thus, the two-syllable vocable "dahin" represents a low-pitched bounce stroke followed by a high-pitched note that is the result of increasing the pressure on the *lunga* ropes. The sound "y" indicates notes produced by decreasing the pressure on the *lunga* ropes. The vocable "diyen" represents a short high-pitch note followed by a long mid-pitched note produced by decreasing the pressure on the *lunga* ropes. Pressure technique sometimes produces a set of three or even four "after-notes". For example, the *lunga* invocation begins with a mid-pitch vocable followed by three notes produced by pressure on the *lunga* ropes; this prolonged sound – mid low high low – requires a four-syllable vocable "deyahiyan".

The onset of notes is often shaped by subtle upward and downward pitch bends. Notes of low and mid pitch may be approached either from above or below, but high-pitched notes only are approached from below. "Neighbor pitches" retain the vowel of the main pitch to which they move over the melodic interval of a major or minor second below or above. Alhaji Abubakari's systematic use of vocables includes this nuance: the vocable "dihin" indicates a high-pitch stroke approached from its lower neighbor; the vocable "deyen" indicates a mid-pitch stroke approached from its upper neighbor.

Three Main Pitches (Vowel)	
-a-	low-pitch tone
-e-	mid-pitch tone
-i-	high-pitch tone
Initial Sound (Consonant)	
d-	normal stroke
z-	first stroke in phrase
Final Sound (Nasal)	
none	short duration
-n	normal duration
Pressure Technique – Increase, Pitch Moves Upward (“h”)	
dahan	from below-low to low
dahen	from low to mid
dehen	from below-mid to mid
dehin	from mid to high
Pressure Technique – Decrease, Pitch Moves Downward (“y”)	
dayan	from above-low to low
deyan	from mid to low
deyen	from above-mid to mid
diyen	from high to mid
Pressure Technique – Moving Multiple Directions	
deyahan	from mid, to below-low, to low
diyahan	from high, to below-low, to low
Press Roll – Multiple Notes From One Stroke	
kr-, gr-	double bounce

Table 1: Vocables for lunga strokes

Examples

I will now illustrate how the *lunga* drum employs these features of construction and instrumental technique to musically render the Dagbani language (see the references for examples of my writing on musical theory, analysis, and criticism of African music).

Praise Name: *Dambobugo*

The answer *lunga* part for *Dambobugo* nicely illustrates the way drumming technique not only helps render the spoken language well, but also makes the musical setting of the language more beautiful (see Figs. 1 and 2).

***Dambobugo* Answer Luŋa Drum Language and Vocables**

Dambobugo din taɣi.	Big family, eat and be satisfied.
To to,	Yes, yes,
Dambobugo din taɣi.	Big family, eat and be satisfied

Vocables	zahan	dayan	dahan	den	deyan*
Dagbani	*dam-*	*bobu-*	*go*	*din*	*taɣi*
Word-for-word	big family			eat	be satisfied

Vocables	dayahan	dayahan
Dagbani	*to*	*to*
Word-for-word	yes	yes

Vocables	zahan	dayan	dahan	den	deyan*
Dagbani	*dam-*	*bobu-*	*go*	*din*	*taɣi*
Word-for-word	big family			eat	be satisfied

My big family, eat and be satisfied.

* Exceptionally fast release of luŋa ropes

Figure 1: *Dambobugo* answer *lunga* – drum language, vocables and translation[6]

6 *Dambobugo* is a musical name for Bukali Kantanparman, a historic chief of Savelugu, a major position in the royal hierarchy of the Dagomba kingdom. After Savelugu-naa Bukali gained the chieftaincy, the weather was particularly good for farming and the harvest was plentiful. The drum language for *Dambobugo* says he has defeated his enemies but continues to think good of everyone. God has rewarded his generous spirit by bringing rain so that big families can eat well and be satisfied.

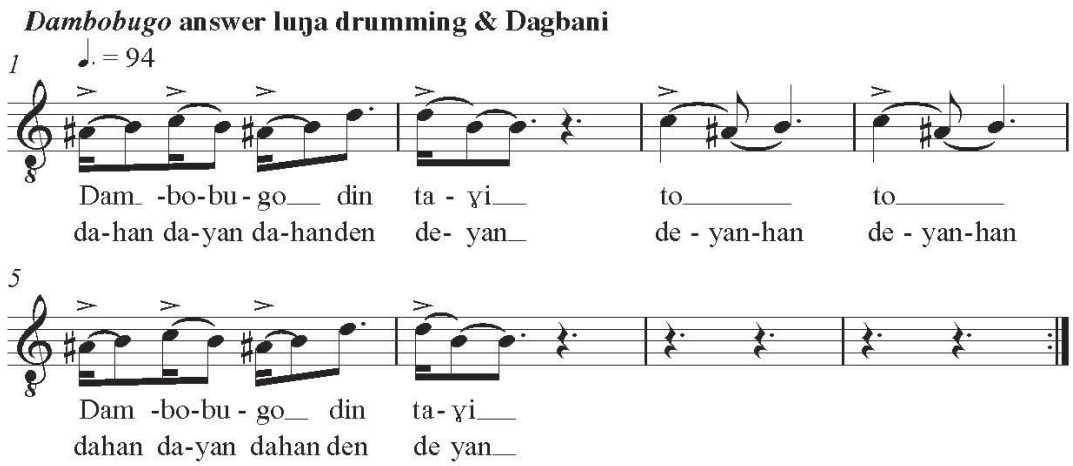


Figure 2: *Dambobugo* answer *lunga* – staff notation

Portamento glides from neighbor pitches below and above the low-pitched main tone enable the answer *lunga* to mimic spoken Dagbani on the word *Dambobugo*, as follows: the nasal ending of the first syllable ("dam-"), the quick timing of the second and third syllables ("-bo-bu-"), and the long duration of the fourth syllable ("-go"). The fourth drum stroke, a simple bounce that is unmodified by after-stroke glides, powerfully presents the all-important verb, 'eat'. Finally, the velar fricative in *tagi* is mimicked on the drum by an unusually fast release of the pressure ropes after the stick's impact with the drum skin set to mid-pitch tone.[7] Before this is repeated, the drum twice sets the word *to* with a note of long duration that gradually fades out over the course of a downward glide to below-low pitch and then an upward scoop to the main low-pitch tone.

The rhythmic idea of the drum strokes in *Dambobugo* is two-in-the-time-of-three within a ternary type metric beat, i.e., 2:3 as two dotted eighth-notes in-the-time-of one dotted quarter-note. The pressure technique not only makes the melody more "curvy", but it adds nuance to the rhythm by subdividing the dotted eighth-notes into two-note pairs. Strongly struck and unmodified by pressure change, the mid-pitch tone made by stroke four stands out in comparison to the "bent" quality of all other notes.

Lunga Invocation: *Nawuni Mali Kpam Pam*

The next example, the ritual invocation that Alhaji Abubakari would always respectfully play whenever he shouldered a *lunga* drum, is much longer (Figs. 3 and 4). Unlike other examples, this music is unmetered and without implicit pulsation.

7 The typography in this manuscript does not use IPA characters; the "g" here represents a velar fricative and the "ng" represents a nasal fricative.

Nawun' Mal' Kpam Pam

Dagbani Text and English Translation

1.	*to to to*	Listen, listen, listen.
2.	*Nawun' mal' kpam pam*	God is powerful.
3.	*To Nawun' mal' kpam pam*	Yes, God is powerful.
4.	*To Nawun' mal' kpam pam*	Yes, God is powerful.
5.	*To Nawun' mal' kpam pam*	Yes, God is powerful.
6.	*Wuninda nam ti dadamnima*	He created human beings.
7.	*N-zah' m-bah' duniya*	He took us and threw us into the world
8.	*To Nawun' mal' kpam pam*	Yes, God is powerful.
9.	*To Nawun' mal' kpam pam*	Yes, God is powerful.

Dagbani, Vocables, and Line-by-line Word-for-word Translation

to	*to*	*to*
deyahin	deyahin	deyanhin
listen	listen	listen

Na-	*wun'*	*mal'*	*kpam*	*pam*	
dan	gra	ziyan	den	den	
God		he has	power	plenty	

to	*Na-*	*wun'*	*mal'*	*kpam*	*pam*
diyan	*dan*	*gra*	*ziyan*	*den*	*den*
yes	*God*		*he has*	*power*	*plenty*

to	*na-*	*wun'*	*mal'*	*kpam*	*pam*
diyan	*dan*	*gra*	*ziyan*	*den*	*den*
yes	*God*		*he has*	*power*	*plenty*

wunin-	*da*	*nam*		*ti*	*da-*	*dam-*	*nima*
den	*den1*	*den*	*deden*	*de*	*diyan*	*den*	*dayin*
he		*created*		*we*	*human beings*		

n-	*zahim*	*m-bahi*	*du-*	*ni-*	*ya*
dan	*ziyan*	*dan*	*den*	*de*	*den*
he	*took [us]*	*he threw [us]*	*[into the]*		*world*

to	*Na-*	*wun'*	*mal'*	*kpam*	*pam*
diyan	*dan*	*gra*	*ziyan*	*den*	*den*
yes	*God*		*he has*	*power*	*plenty*

to	*Na-*	*wun'*	*mal'*	*kpam*	*pam*
diyan	*dan*	*gra*	*ziyan*	*den*	*den*
yes	*God*		*he has*	*power*	*plenty*

[1] slightly higher pitch than previous "den"

Figure 3: *Naawuni Mali Kpam Pam* – drum language, vocables and translation

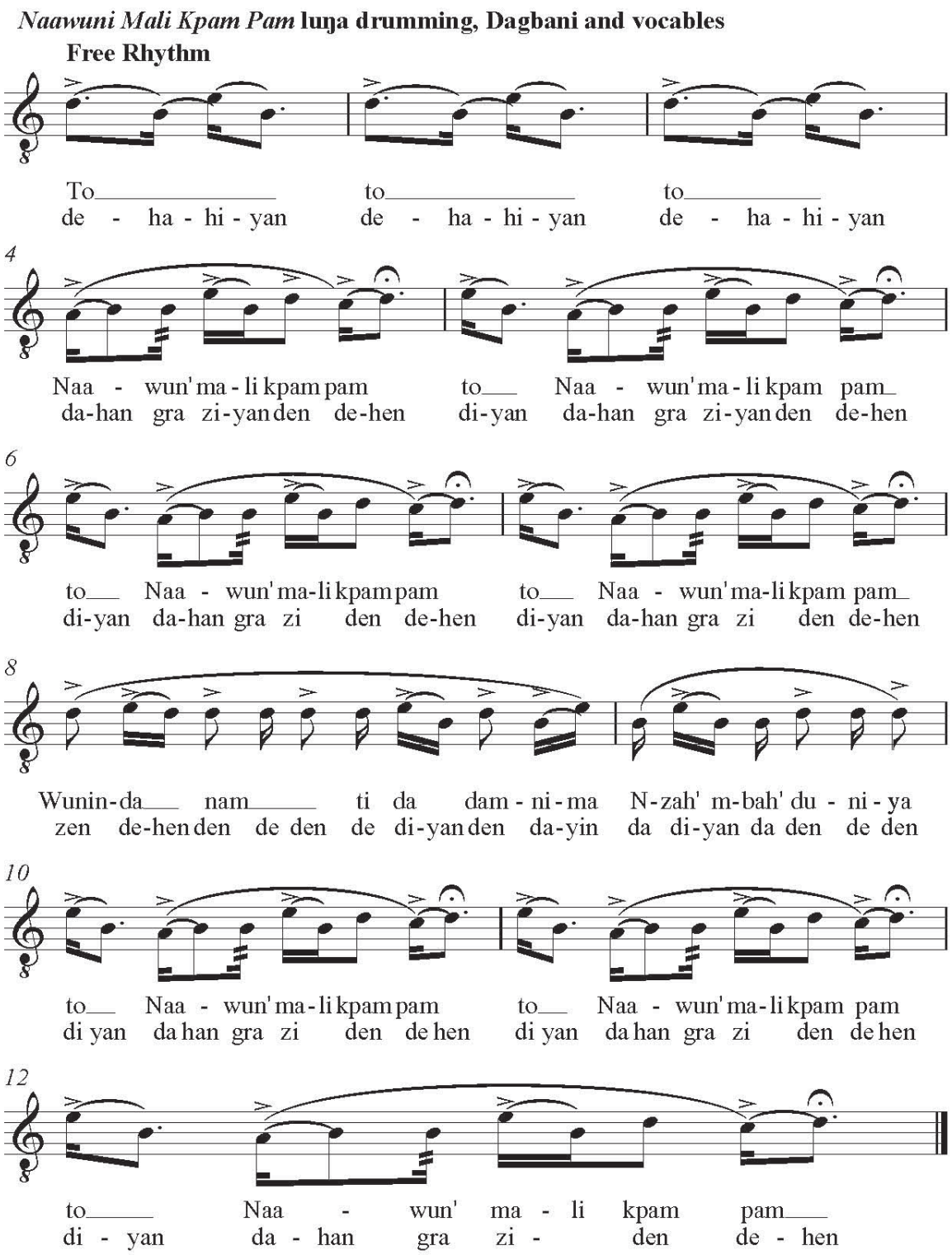


Figure 4: *Naawuni Mali Kpam Pam* – staff notation

Alhaji's way of chanting the drumming vocables is virtually identical to his drummed version so the comments below apply equally well to both (see Table 2).

line 1	The multiple pressure notes mimic the slow decay of the final vowel.
line 2	The upward *portamento* mimics the subtle upward glide of the "Na-" syllable.
line 2	The turned-stick press roll mimics the disyllabic "-wuni", whose final vowel is often not articulated.
line 2	The downward glide mimics the disyllabic *mali*, whose final vowel is often not articulated.
line 2	The straight-stick stroke renders *kpam*, but an upward *portamento* mimics the subtle upward glide into the syllable *pam* from the final "m" of the preceding word. The contrasting onsets of the two mid-pitch notes are pleasing.
line 3	The downward glide mimics falling tone and slow decay of the word *to*.
lines 4–6	Repeat line 3.
line 6	A trisyllabic word, *wuninda*, is rendered with two stick strokes and a downward pressure glide *zen deyen*.
line 6	The drum repeats two syllables of "ti da-" for purely musical effect (see Euba 1990: 235 on repetition).
line 6	The downward glide mimics the longer duration of "da-"; the longer time value similarly mimics the longer duration of "-dam-"; the upward glide mimics the disyllabic structure and upward intonation of "-nima".
line 7	The short time value of the low-pitch first stroke and the downward glide after the second drum stroke mimic the rhythm, intonation, and syllable-ending sound of "N-zahi", which is elided with the nasal beginning of "m-bahi" ("m-" is the pronoun 'he').
line 7	Four simple bounce strokes render the four syllables "-bah' du-ni-ya".
lines 8–9	Repeat line 3

Table 2: *Nawuni Mali Kpam Pam* – comments

Praise Name: *Nagbiegu*

Nagbiegu is the appellation for Naa Abudu. Abudu's predecessor had taken land from a neighboring polity whose chief was threatening to take it back (see Fig. 5). Naa Abudu stood firm, saying that he is a "bad cow" and challenging his enemy, "Try me and see!" The drummed praise name is in a two-part verse-chorus form. In the verse section, the lead *lunga* plays phrases of drum language that are punctuated by single booming notes on *gung-gong* and answer *lunga* (Fig. 6). Upon the lead *lunga's* signal the answering drums sound the distinctive theme of the composition.

***Naɣbiɛgu* Lead Luŋa Drum Language**

Call

To to to,	Listen,
Chɛguru, Chɛguru, Chɛguru,	One who never runs away,
Doo bi nyɛ o to dapala,	Son of who never runs away,
Abulaibila to Chɛguru.	Abulai, yes, the one who never runs away.

chɛ	refuse
guru	run
Abulaibila	Abulai, the younger

Signal to play theme

To Naɣbiɛɣu ŋuni daa ku o.	Yes, Naɣbiɛgu, he killed him.

naɣbiɛɣu	bad cow
ŋuni	he who
daa	v. past

Talk 1

Chɛguru, gbam o, Chɛguru.	Chɛguru, catch him, Chɛguru.

or

Chɛguru, kum o, Chɛguru.	Chɛguru, kill him, Chɛguru.

gbam o	catch him
kum o	kill him

Talk 2

Chɛguru, Chɛguru, Chɛguru, gbam o.	Chɛguru, grab him.

Figure 5: *Nagbiegu* – drum language, vocables and translation

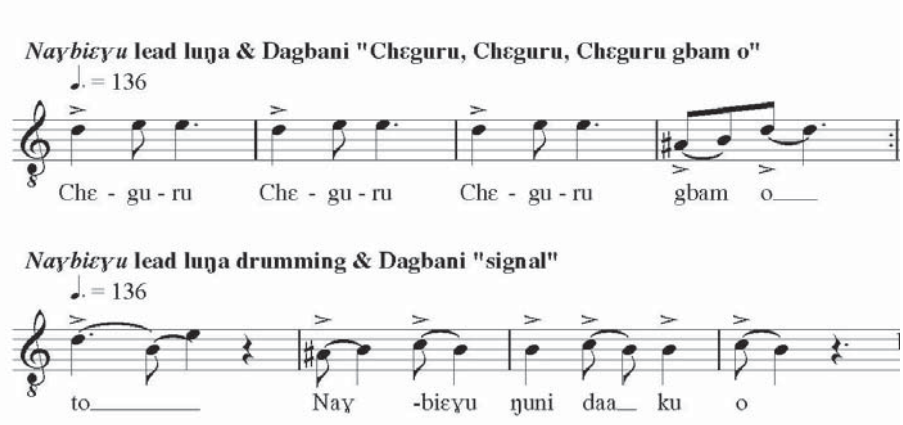


Figure 6: *Nagbiegu* – staff notation

I have selected for attention a verse phrase that aligns the drum's contrast in the loudness of strokes ("DEN didin") with the stress pattern in the spoken Dagbani – *CHE-guru*. The drum's rhythm also matches the duration of these syllables. This example once again illustrates the very close mimesis between language, vocables, and drumming. The signal phrase, however, enables us to consider a case where the drumming does not perfectly match the language.

The question is why the two-syllable word *nguni* is rendered with a simple note, uninflected by pressure bends. I hypothesize a musical reason. The rhythm of the signal opens with accentuation of the dotted quarter-note beats (mm. 1–2) but then shifts to accentuation of a cross rhythmic quarter-note feel, which is also a prominent feature of the chorus theme (three-in-the-time-of-two, or 3:2). The rhythm of the final four drum strokes in the signal achieves multidimensional depth by virtue of the similarity between the "straight" strokes one and three and the "bent" strokes two and four.[8]

Conclusion

This paper illustrates the remarkably adept manner in which the Dagomba *lunga* drum can mimic the sound of spoken Dagbani, whose texts underlie much of its drumming. In addition to the correlation of musical pitch to lexical speech tone, the *lunga* can imitate many subtle phonetic elements. From the perspective of the music itself, the drumming becomes sophisticated in its rhythm, subtle in its melody, and clear in the shape of its phrases when phonological features of the language are mimetically represented. In a word, the music becomes beautiful.

8 An anonymous reviewer notes a phonetic reason as well: two-syllable "ngu-ni" can be shortened to the mono-syllabic "ngun'".

References

Agawu, Kofi. 1995. *African Rhythm: A Northern Ewe Perspective*. Cambridge etc.: Cambridge University Press.

Euba, Akin. 1990. *Yorùbá Drumming: The Dùndún Tradition*. (Bayreuth African Studies Series, 21/22). Bayreuth: Breitinger.

Fiagbedzi, Nissio. 2005. An Essay on the Nature of the Aesthetic in the African Musical Arts. Lightning Source.

Fusheini, Hudu. 2007. "Morphological category and nasal contrast in Dagbani". In: Carter, Nicole et al. (eds.). *Proceedings of the 22nd NorthWest Linguistics Conference, February 18–19, 2006*. Burnaby, B.C., Canada: Simon Fraser University Linguistics Graduate Student Association, 119–132.

Hyman, L. 1993. "Structure preservation and postlexical tonology in Dagbani". *Phonetics and Phonology* 4: 235–254.

Locke, David. 1990. *Drum Damba: talking drum lessons*. Crown Point, Ind.: White Cliffs Media.

Locke, David. 2010. *Dagomba Dance-Drumming*. Online: <https://wikis.uit.tufts.edu/confluence/display/DagombaDanceDrumming/Welcome>.

Olawsky, Knut J. 1996. *An Introduction to Dagbani Phonology*. (Theorie des Lexikons, 76). Düsseldorf: SFB 282 (Heinrich-Heine-Universität).

Olawsky, Knut J. 1999. *Aspects of Dagbani Grammar*. München: Lincom Europa.

Villepastour, Amanda. 2010. *Ancient Text Messages of the Yorùbá Bàtá Drum: Cracking the Code*. Farnham: Ashgate.

Wilson, W.A.A. & J.T. Bendor-Samuel. 1969. "The phonology of the nominal in Dagbani". *Linguistics* 52: 56–82.

The Relationship between Language and Music in the Bakonzo Culture (Uganda)

(Vanna Viola Crupi)

Introduction

This paper shows some results of a preliminary study of the relationship between language and music of the Bakonzo, a Bantu-speaking people of Uganda[1]. The main aim of this work is to put forward some hypotheses about the influence of linguistic tones on the melodic profile of Konzo songs, to analyze the connection between linguistic and musical "downdrift" and to observe some effects of Bantu "alliterative concord" on Konzo lyrics. Bakonzo culture (Pennacini & Wittenberg 2008) is particularly suitable for the aims of this research due to the characteristics of both the linguistic and musical systems of the Bakonzo people. The culture of the Bakonzo people is distinguished in various ways from the cultural model of the Great Lakes area where they live. The Bakonzo language (*lhukonzo*), for example, differs from many of the Bantu languages spoken in the territory (Ladefoged, Glick & Criper 1972). In the same way, their pitch system can be defined as "pseudo-heptatonic" and differs from others in the Great Lakes area, where one finds mostly pentatonic systems. This makes this culture particularly suitable for the aim of the analysis (Crupi 2008; Facci & Nannyonga 2008). Indeed, the geographical and cultural isolation of Bakonzo seems to have saved them from external cultural influences, possibly for a longer period than other populations of the area. This work adopts some experimental approaches in methodology, using tools and strategies borrowed from the disciplines of phonetics and cognitive psychology.

Previous studies and future directions

1. Ethnomusicology

As for the relationship between music and language, the most important scientific results obtained so far by ethnomusicology concern, on the one hand, the reproduction of coded verbal messages by means of sound-producing instruments and, on the other, the relationship between song melodies and specific tonal languages. In the first case, in-depth research

1 The ethnographic research was conducted with the valuable contribution of the doctoral program in History and Analysis of Musical Culture at the University Sapienza of Rome and that of the Ethnological Mission in Equatorial Africa having supported my research in Uganda since 2005.

has been devoted to the strategies involved in the production and reception of whistled and drummed messages on the African continent.[2]

In the second case, representing the framework of this research work, there is strong evidence for a close relationship between melody and language tones in the sung repertoires of many cultures of Africa; the hypothesis appears to be widely shared by many Africanist musicologists: Agawu (2001, 2003) Arom (1985), Blacking (1973), Cooke (1980), Facci (1996), Gray (1995), Jones (1959), Kubik (1994, 1999, 2010), Nannyonga (2007), Nketia (1958, 1975). All these scholars, although with different approaches and levels of detail, dealt with problems related to the complexity of such a relationship. The wide experience gained in African cultures by the Austrian scholar Gerhard Kubik led him to write that African music is so closely connected to language "that it is hardly possible today to study it without the necessary background in African languages" (Kubik 1994: 9).

John Blacking's considerations of Venda music deserve special mention. This scholar deals with the topic of the relationship between linguistic tones and melodies – analyzing two children's songs, from which he draws some interesting conclusions (Blacking 1973: 69–70):

> [...] Variations in melody and rhythm may therefore indicate not musical preferences, but the incidental consequences of changes in speech tone, which are themselves generated by the use of different words [...]. This does not mean that the Venda are unmoved by music, or that they regard it as a mere extension of language. [...] Even in solo vocal music like the children's songs, the form of melodies can be divided into call and response sections [...] It is only in the call section of the songs that melodies follow the speech-tone patterns of words, and also the general rule that each syllable of a word may be accompanied by only one tone.

John Blacking recognizes then that the responsorial Venda melodies follow the patterns of speech intonation only in the first part of the song: the utterances of the soloist. The melodic autonomy of choral parts from linguistic rules seems to lie in the natural repetition of the choral responses. Applying Blacking's analysis to the Konzo songs, one finds that this melodic autonomy tends to affect even the solo parts. Typically, this happens only when words or verses have already been sung at least once in respect of the prosodic profile. A possible interpretation would be that there is no need for making the lyrics intelligible – they are repeated and already known; rather, this kind of performance gives the soloist a wider margin of melodic freedom.

2 Cf. Arom (1985), Alexandre (1969), Carrington (1949a, 1949b), Euba (1990), Kubik (1994, 2010), Rouget (1964), Sebeok & Umiker-Sebeok (1976), Stern (1957), Zemp & Kaufman (1969), Villepastour (2010).

The hypothetical rule: linguistic repetition > melodic autonomy may be better understood by considering the studies of musicologist Kofi Agawu, one of the scholars who worked more on the relationship between linguistic tone and melody in Africa. Agawu analyzes all variations of a melodic phrase (*ayaa do*) within an Ewe song characterized by the sequence of tones HL (High – Low), concluding that a melodic distance from the prosodic profile of the lyrics in an Ewe song always appears after a version that closely respects the pattern of tones (thus preserving the intelligibility of words).

The author specifies that "the repeated phrase *ayaa do* was properly heard before each musically 'wrong' intonation" (Agawu 2001: 949; translation VC). In addition, Kofi Agawu, while sharing Kubik's assumption, considers the relationship between linguistic tones and melodies not as a univocal effect of language on melody. He draws the conclusion that "the tone of the language can affect the melody but cannot determine its structure" (ibid.).

Starting from this assumption this paper will examine the relationship between linguistic tone and melody in Konzo songs, with the help of some experimental analyses, taking into account the influence of specific rules of Bantu languages on the sung melodies.

2. Neuroscience

In many Bantu-speaking cultures of Africa, linguistic tones are neither stable nor absolute. Even considering only the level tone languages, excluding the languages that use contour tones, most of them are subjected to a downdrift process, a characteristic that, according to neuropsychologist Aniruddh Patel (2008), makes difficult any comparison with the system of pitches used in music. This scholar considers the "discrete level tone languages" – a category of level tone languages "not subject to the downtrend" – as "the best languages for comparing pitch contrast in language and music" (Patel 2008: 43). On the other hand, Patel (2008: 44) argues that "even if such languages do not show declination, it seems very likely that discrete level tone languages do not use fixed pitch intervals." In this paper I try to suggest that the phenomenon of downdrift does not completely prevent the comparison between music and language, rather might even encourage it. In an analysis of vocal Konzo music the falling melodic contour of the songs could in fact be traced back to the linguistic phenomenon of downdrift, as will be shown later.

This may confirm our view that a comprehensive and extensive study on the relationship between music and language cannot do without due consideration of the different categories of tonal languages which are widespread on the African continent.

Analysis of the relationship between melodic and prosodic profile in Konzo songs

1. Methodology

The methodology adopted in this research work brings together analytical tools borrowed from different disciplines, such as experimental phonetics, musicology and cognitive science. Given the absence of detailed studies[3] on the lhukonzo language[4] (Tucker 1960; Ladefoged, Glick & Criper 1972), it was necessary to verify, during the preliminary phase of my research, whether this language could be considered in all respects a tonal language, confirming the existence of linguistic and semantic relevance of tones. Once I had identified a sufficient number of tonal minimal pairs, I started with a first perceptual analysis to assess the ability of detection of tonal differences by the Bakonzo people as well as their understanding of the different meaning of word pairs in an absence of sentence context. Another experiment involved the comparison between the melodic contour of singing and declamation of related texts. A third experiment in perception, based on the improvisation of songs, was conducted to determine the way musicians would recognize the differences in linguistic tones and respect them in their vocal performances. This test, which involved musicians of various ages and levels of education, will be briefly presented in the last part of this paper. It has provided useful data to support the hypothesis of an influence of language on melodies of songs and to determine the high level of musicians' perception in recognizing the tonal variations in language.

2. Main features of Konzo vocal music

The traditional repertoire of Konzo songs can be categorized in songs for birth (songs for the gestation, songs for child birth, naming songs, lullabies, songs for twins), songs for initiation (*Erikenza*: female initiation songs, *Olhusumba*: male initiation songs), wedding songs (songs for the bride, songs for the Wedding Party), funeral songs (general funeral songs, lament songs for initiated men, songs for women and uninitiated men), religious

3 When I started my research in 2007, more studies were available on the language of Banande, people who share with the Bakonzo the same origins and culture but live in the Congo area and speak as their second language French instead of English. While the work was going on, a Nande/Konzo bilingual dictionary was published; see Mutaka (2008).

4 In 2008 I had the opportunity to meet the Italian linguist Barbara Turchetta, who greatly contributed to my research, working with me in the field in Uganda. She is still carrying out some important research on the lhukonzo language. Turchetta kindly

songs (liturgical music, songs for the purification of the ridges,[5] worship of *Kubandwa* spirits); political songs (political hymns and protest songs).

From the analysis of the collected repertoires some recurring characteristics of the traditional Konzo songs clearly emerged. Many of them indeed fall into a dominant model characterized by the following features:

- bipartite formal structure (mostly);
- responsorial form (mostly) solo / chorus;
- variable metrical pattern (in different songs);
- cyclic flow (rhythmic patterns are repeated with or without variations);
- generally falling or undulating melodic contour;
- syllabic melodies;
- scalar models mainly hexatonic or heptatonic and pentachordal (these latter seem more prevalent in children's repertoires and modern tunes);
- melodic variation (which generally may occur when the singer improvises new verses in a song or when he sings the same words in repeated verses);
- overlapping voices unsystematic (that between solo voice and chorus if present is generally of short duration).

We can assume however that each Konzo song has its position on a hypothetical continuous line from speech to singing: rhythmical and emphasized speech, songs without stable pitch, songs with formalized melody and rhythm. It is not uncommon, moreover, that the Bakonzo improvise on their own songs, using rhythmic and melodic patterns as formulas. On these rhythmic-melodic reference grids, the singers seem to compose new verses.

3. Phonetic analysis

I will present a concrete example of profile comparison between prosodic and melodic contour of a Konzo song (*Byana byanyonyi*) using the method of phonetic analysis. The same experiment was done with several songs giving in many cases similar results[6]. In order to perform the experiment the text of the recorded songs was first transcribed.

5 The traditional administration of the Konzo territory (Gardoncini 2009) is based on a subdivision of the mountain ridges and sub-ridges. Each ridge begins high up in the forest and is delimited for its entire length by two rivers, finishing at their convergence. Every ridge (obulhambo) has a chief, the omukuru w'obulhambo, who, when in the past there were troubles in his territory, performed a purification ritual called eribhiria obulhambo. It includes musical performances both in the forest and in the inhabited area, even though the two performances take on different features. In this complex ritual that takes place mostly in the mountains of the massif, the songs have a key role. In specific sacred places the men sing accompanied by the sound of the endara xylophone (Crupi 2013).

6 For more examples see Crupi (2011).

As a second step I recorded the declamation of the song texts performed by lhukonzo speakers.[7] Finally I compared the original sung version with the declamation of the lyrics.

The responsorial song Byana Byanyonyi is characterized by a cyclical form (solo + chorus = 1 cycle) in which the part of the chorus remains the same while the text of the soloist varies. All verses can, however, be traced to a single heptatonic model, though the pitches may fluctuate. The rhythmic accompaniment based on triplets is usually performed by a drum; for purposes of the analysis the proposed version was recorded even without drum accompaniment. This song was sung in Nyakalenghigya village by an old woman, Regina Masika, a birth attendant who speaks Lhukonzo and Rutoro, and two other women, her daughters Mary (English-speaking) and Josephine. [8] The response of the chorus is the same in all verses: "*Tumbelele tumbeleleko byana byanyonyi*", while the text of the soloist changes.

For the sake of brevity, I chose to present only the analysis of a single verse of the selected song. The choice fell on verse no. 6 for the presence in it of a word which is part of a tonal minimal pair (*byasekèra/byasekéra* – 'they laugh'/'they weed').

The aims of the phonetic experiment were to detect the interval between adjacent tones in two versions (spoken and sung) of the same verse, to identify the degree of proximity between the prosodic profile of the song and the melody line.[9] For this type of analysis, by convention, I indicate in the texts tone H (High) and L (Low), and even the tones which are neither considered low nor high and marked by the letter M (Medium) as shown in the following subjective transcription.[10]

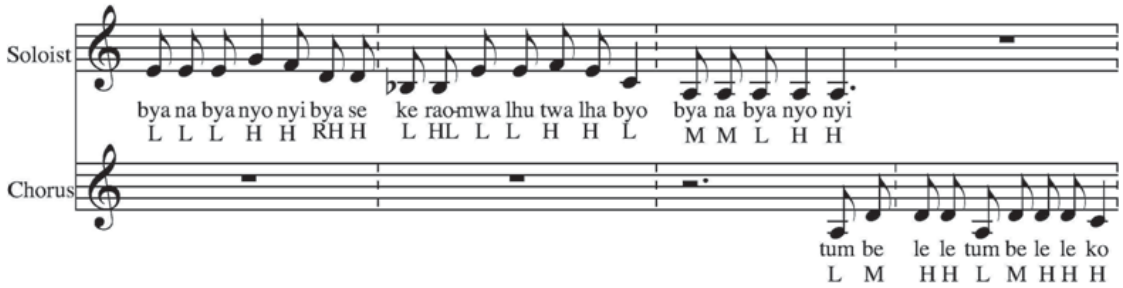
---

7 The request made to the performers to recite the words of the song was a complete failure because their memory of the sung texts was closely linked to the related melodies. For a good 70% of the singers it was therefore difficult to separate the words of the song from the melody. After many attempts I have opted for an alternative solution: the transcription of the texts and the subsequent recitation of the same by Justine Muhindo, a speaker of Konzo and English who collaborated on the research as interpreter. The declamation of a second speaker, Florence, a primary school teacher of Ibanda, helped me to have a control parameter for the experiment.

8 Song recorded on 18 August 2008 in Nyakalenghigya village (Kasese district).

9 For reasons of space it is not possible here to provide a comprehensive musical analysis of the song. For a broader analysis of this and other Konzo songs see Crupi (2011).

10 Each tone is assigned to a specific tone class, considering its relationship with adjacent tones. Regarding the tone classes, Tucker refers to three categories: H, M, L, whereas the work of Mutaka on Kinande identifies only the tones H and L as distinctive. In Barbara Turchetta's three tone analysis the category of middle tones is assimilated with the absence of tone (Turchetta, in print). The presence of ascending and descending distinctive tones (tones giving a distinction in meaning) is still to be proved in the Konzo language.



Score 1. subjective transcription of verse 6 of the song Byana Byanonyi. The song was sung by Josephine (soloist), Mary and Regina Masika (chorus) in Nyakalenghigya village. Sung version without drum accompaniment. The letters H, M and L written below the lyrics show the tones of the declaimed text measured with the phonetic analysis program Praat.

It will be useful to give an example of a phonetic transcription of the sung and recited version of the selected verse to highlight the variations in pronunciation of individual words between the two utterances.

Solo part – verse 6

Speech:	Bjana	βjaɲoːɲi	Bjasekɛːra	omwaɹuthwəɹwaːbjo
Singing:	Bjaɲːa	βjaɲɔːɲi	Bjasekeːr’	omuɹuthwɹabjo

Translation: Little birds, they are laughing in their nest

The syllabic elision occurs for syllables showing little relevance on the level of meaning. All that is elided is, indeed, the final a, which is a default vowel in the inflectional system. You will notice in the following graphs that the frequency of the sung voice appears higher than the speech.

Byana Byanonyi – Phonetic analysis
Verse 6 – speech

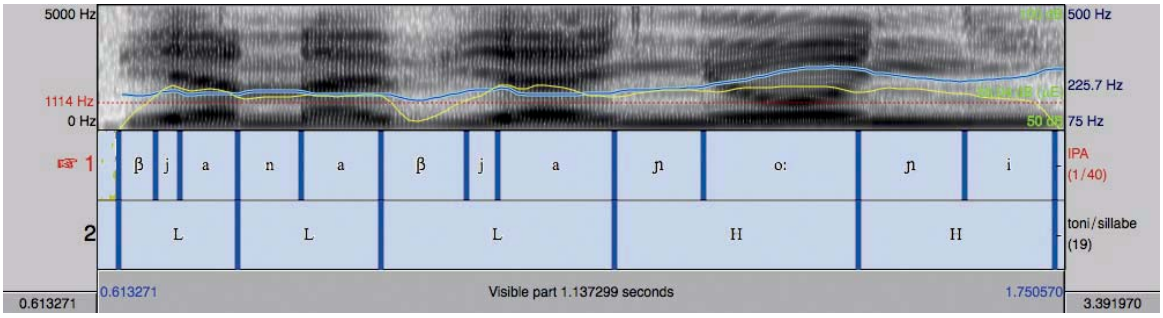


Figure 1a

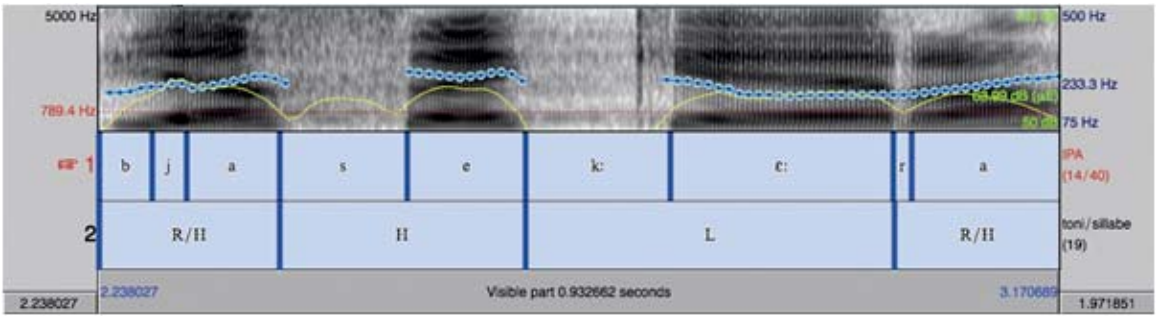


Figure 1b

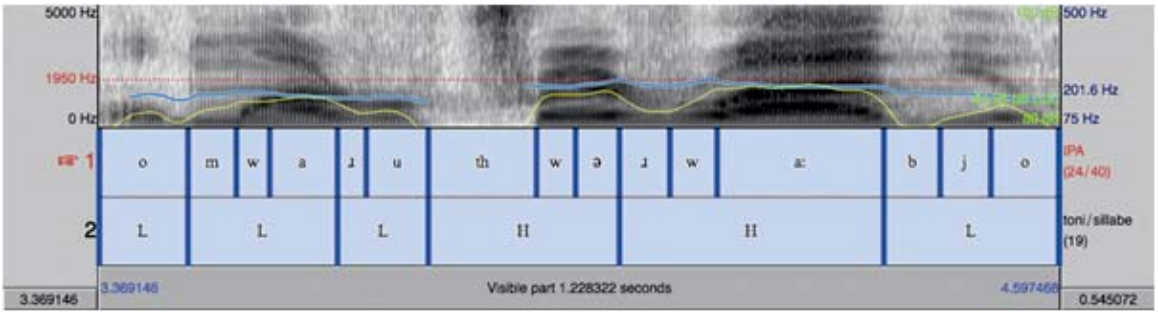


Figure 1c

Figure 1a-c: Phonetic analysis of verse V6 of the song *Byana byanyonyi*. The graphs include: spectrogram, phonetic transcription of the text, tone language. The blue line indicates the pitch contour while the yellow line represents the intensity. Analysis program: Praat.

Verse 6 – singing

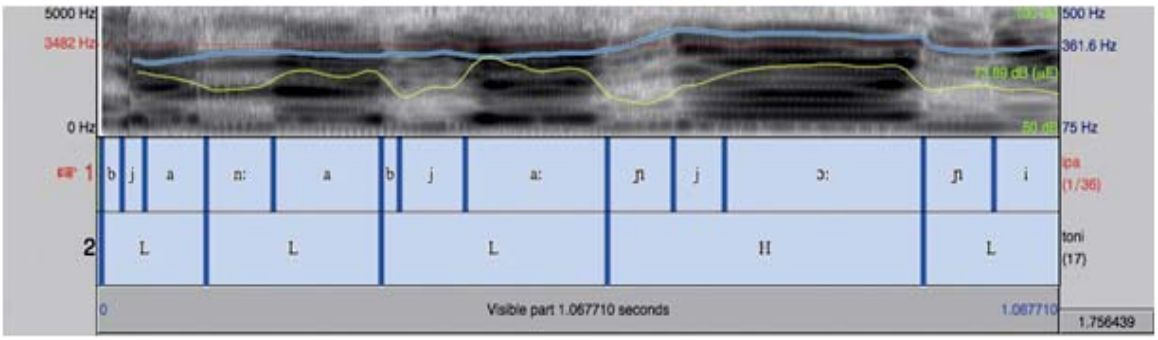


Figure 2a

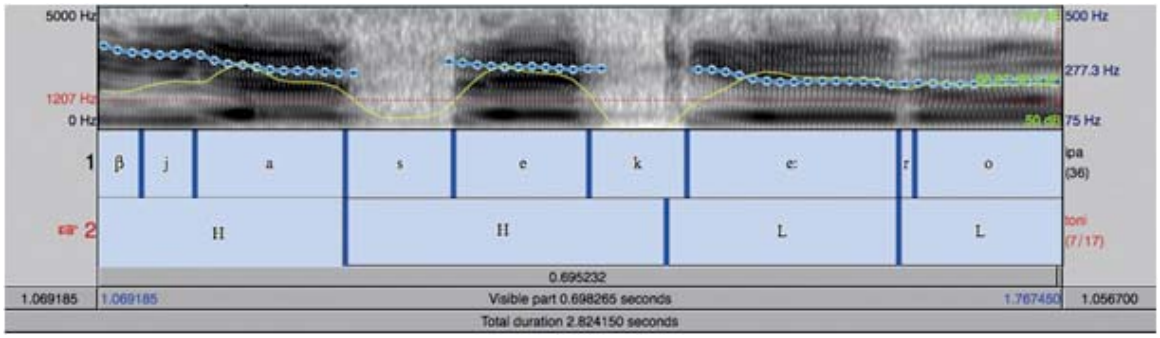


Figure 2b

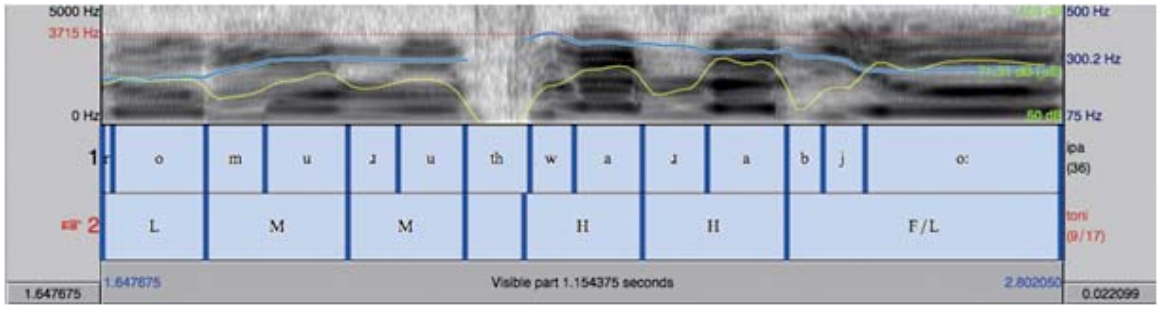


Figure 2c

Figure 2a-c: Phonetic analysis of the song *Byana byanyony*i (verse 6). The graphs include: spectrogram, phonetic transcription of the text, linguistic tones. The blue line indicates the pitch contour while the yellow line represents the intensity. Analysis program: Praat.

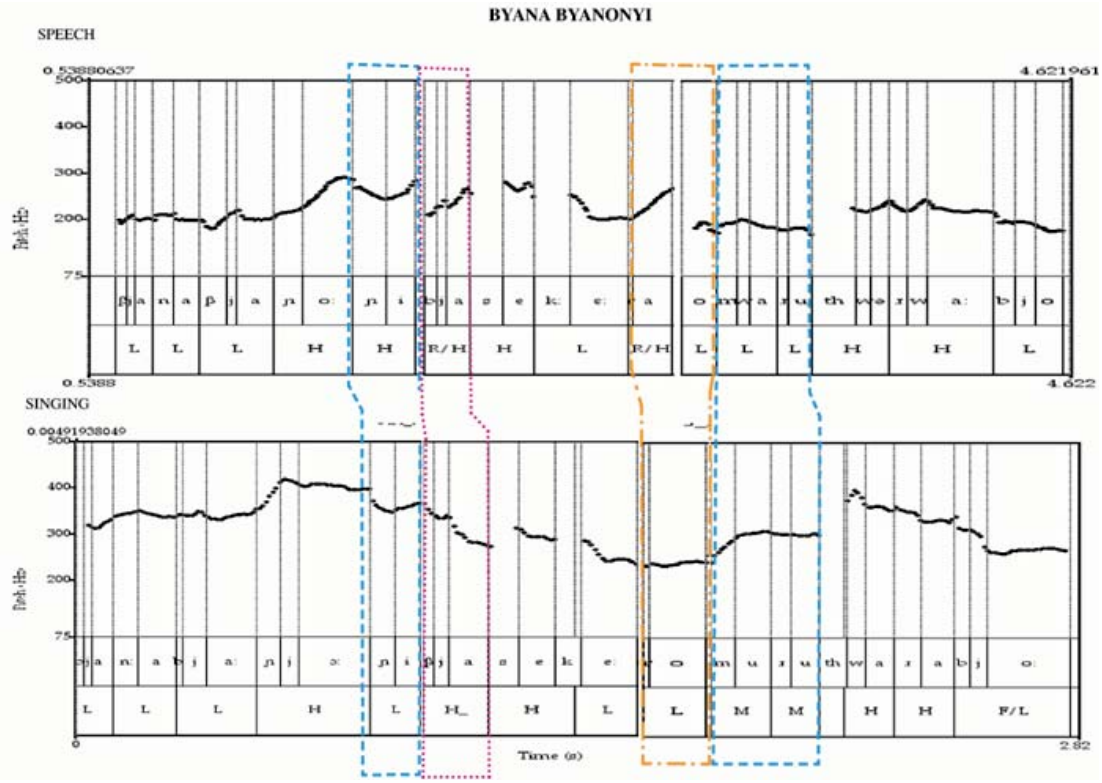


Figure 3. Comparative Chart: profile comparison between prosodic and melodic contour of the first verse of the song *Byana byanyonyi*. The vertical dashed lines highlight elements of discrepancy in the two profiles. The ordinate indicates the pitch in Hertz, the abscissa the time. Analysis program: Praat.

In this as in the previous graphs, the unit reference considered for the comparative analysis is the syllable. Each syllable is characterized by its tone (H, L, M). On the last syllable of each spoken word a rising intonation is noticeable, but this is not always a feature when singing – probably because of my request to clearly articulate the text declaimed, which caused the speaker [11] to focus on each word in isolation. This does not apply to the last word of the entire sentence, which shows a falling contour. It is important to note that in all the words of this sentence the pertinent tone falls on the penultimate syllable.[12] Moreover, both in our case and in many others, the melody of the song conforms to it. In reading the given data, it is plausible to state that melodies, in general, respect the pitch of pertinent tone of each word (providing semantic opposition).

As analyzed in the sample with its few variations between sung and recited text, there are many other cases in which partially sung words undergo changes with respect to proper pronunciation. Usually this happens because of the need to adapt the text to the metre of the song. Still, these 'deformed' or modified words remain somewhat understandable due to the context of the sentence or the topic of the specific repertoire to which they belong.

4. Downdrift and melody

In the example discussed above, the prosodic profile of the utterances, rising falling, shows a general downward drift, which is almost the same in speech and singing. This takes into account a phenomenon detectable in all tonal languages defined by linguists as downdrift[13] (Nespor 1994: 99):

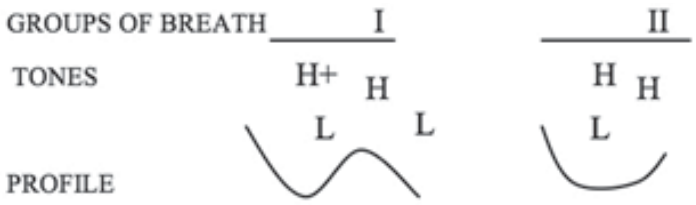
> [...] In many African tongues there is a phenomenon which we will call tonal descending derives (downdrift / downtrend) that consists in a gradual descent of the height of identical tones in phonological terms and belonging to the same phonological constituent corresponding to a breath group[14] [...].

11 Justine Muhindo, a speaker of Lhukonzo and English, is a guide at the Ruwenzory National Park; she lives in Ibanda village (Kasese District).

12 In Lhukonzo as in Kinande the pertinent tone generally falls on the penultimate syllable. From a perceptive point of view, the penultimate syllable is relevant for Lhukonzo speakers because it is the syllable where tones offer meaning opposition in the same phonetic sequence.

13 See also Patel (2007: 43) and Cardona (1988).

14 In phonetics, a sequence of sounds articulated in the course of a single exhalation is called a breath group; it is an utterance or part of an utterance that is produced between pauses for breath.



This means that within a "breath group" in a sequence of tones HLHL, the pitch of the second tone (H) will be lower in pitch than the first, and consequently also the second tone low (L) will be lower than the previous one. This downdrift generally applies only to each breath unit. In the presence of two or more units (or breath groups), a sentence will have an internal undulating prosodic pattern:

Despite this general rule, the analysis makes clear that the higher tones in a sentence are often those in the first breath group.[15] In the first graph, as is shown in the previous paragraph, there is in fact a clear minimal and progressive lowering of the pitches, both in speech and in singing. It follows that the downward trend of the sung melodies in Konzo songs can be considered, at least in part, as a direct derivation of the phenomenon of downdrift.

5. Alliterative concord

Another phenomenon related to the musicality of speech that could be detected in Lhukonzo and in general in Bantu languages is alliterative concord, rarely evaluated in the analysis of sung verses. This linguistic phenomenon has been well described by the linguist Funwi Ayuninjam in his work on the Mbili language (Ayuninjam 1998: 20):

> Bantu languages are characterized by a system of noun classes, variously referred to as 'grammatical class gender', 'nouns class concord', or 'alliterative concord', as the noun system often calls for a repetition of the same syllable as a prefix of each successive word in the phrase, or as concordial coefficient in all associated parts of speech. The alliteration is therefore semantic.[...] The prefix to be used depends on the part of speech of the form (noun, adjective, verb, etc.) or its number (singular or plural). In some cases, a class prefix composed of a cv sequence may have as its coreferential element the consonant alone excluding the vowel.

Alliterative concord consists in the repetition of syllables of a single prefix in many words in a sentence. I noted during my studies that for the agglutinative characteristics of Lhukonzo, the phenomenon of alliterative

15 The musicologist Laz Ekwueme argued in reference to Bantu music: "When the initial syllable of the first word or phrase is a high tone, the melody starts with the highest note in the whole tune." (Ekwueme 1974: 48).

concord could determine or at least facilitate the presence of assonances in proverbs, telling stories and songs. In many cases, the similarities in the individual sentences are increased by the continued repetition of whole words. Alliterative concord is also joined by the phenomenon of 'reduplication', which consists in doubling the root of the verb to add semantic meaning and enhance intensity or continuity of the given action (e. g. *erítuma* 'send', *erítumatúma* 'send with urgency'; *erihúma* 'beating', *erihumahúma* 'beating rapidly'). This phenomenon creates the same perceptual effect of alliteration. Both linguistic phenomena give rise to the proliferation of morphological similarities between contiguous words in the context of the sentence and seem to have a role in the presence of the many assonances found in Konzo lyrics. The same *Byana byanyonyi* song also gives a valid example of alliterative concord (prefix bya-). Although the primary function of alliteration is grammatical, it does not exclude the Bakonzo from using them for poetic purposes. It will be difficult, however, in the presence of alliteration, to understand whether they are the result of metrical poetic rules or are already originally present in the spoken language. The role of alliterative concord in the songs appears important not only in the study of the metre of sung verses but also in the analysis of their melodic profile, because the repetition of the same syllable in a verse also involves the repetition of linguistic tones that characterize the syllable itself; therefore it can affect the whole prosodic profile of the verses. In the light of these considerations, alliterative concord should be regarded as one of the factors particularly useful in the study of the relationship between music and language in Bantu songs.

Until now I have presented some analyses based on phonetics. As I said, part of my field research was devoted to the perception of linguistic tone. The use of a methodology borrowed from cognitive disciplines was necessary not only to define the degree of perception and recognition of tones in speech, but also to identify their perception in songs. For this purpose I developed specific perceptual experiments: an execution test (for singers and musicians) and a test on understanding the semantic content of songs.[16]

16 The first experiment was made on the basis of guidelines provided in the initial planning phase of the research activities by members of the laboratory "Cognitive psychology and cognitive science – psychology of sound, rhythm and music", Rome, coordinated by Professor Marta Olivetti Belardinelli. The second experiment was planned in collaboration with the Italian linguist Barbara Turchetta, the phonetician Laura Mori and the valuable support of Augustine, a young Konzo musician, informant and interpreter during the last two missions of research, who has cooperated in the creation of sound examples.

6. Perception and analysis: vocal and instrumental improvising on tonal minimal pairs

For the perception experiment singers and musicians of various level of experience were involved. They were asked to listen and explain the meaning of 15 pairs of sentences characterized by tonal minimal pairs. The first aim was to check their understanding of the pairs of sentences.

Then I asked them to choose one of the 15 pairs of sentences and improvise a song with it. Finally, the correctness of linguistic tones according to language rules during vocal or instrumental improvisation was verified. In all cases, the musicians were very familiar with identifying the sound differences between the tone pairs; at the same time they could not always recognize the difference in meaning of the sentences.

During the improvisation, the singers sometimes found it difficult to create a song and instead sang the text as rhythmical and emphasized speech. In almost all cases in which the melody used was stable, I discovered that the musicians borrowed it from traditional songs, which were adaptable to the metre of the chosen sentences and their linguistic tones. Once the melody fitted, the differences between tonal minimal pairs in the song were expressed respecting the movements of tones of speech. Even considering ambiguous cases, probably due to the presence of tone and accent in the same position, it was possible to detect differences between two sung words.

Rarely were the pairs of sentences on which the experiment was based sung in isolation. Some musicians used their instruments (flute or harp) to accompany their performances. In the case of Erissanya,[17] an expert *enyamulera* (flute) player living in Kiondo village (Kasese district), he chose the following pair of sentences:

A Esyóngwé: sisiri hano[18]	B *Esyòngwé sisiri hano*
A Firewood is not here	B Leopards are not here

The songs accompanied by *enyamulera* usually consist of alternating instrumental solo parts and vocal parts.

Erissanya began to improvise with verses of his own invention, alternating with the sound of the flute; in the same song, he used the two sentences with the minimal pairs A and B while he varied the melodic singing at the two target words: *esyóngwé* and *esyòngwé*. It was not possible to verify if Erissanya borrowed the melody from a song of his repertoire, adapting the new pair of sentences to it; in any case, as he said, he was composing a new song with the proposed text.

17 Interview conducted on 14 August 2007 in Kiondo village (Kasese district).

18 In the first phrase (A), the first syllable of hano has a high falling tone, in the second phrase (B) bears a low tone.

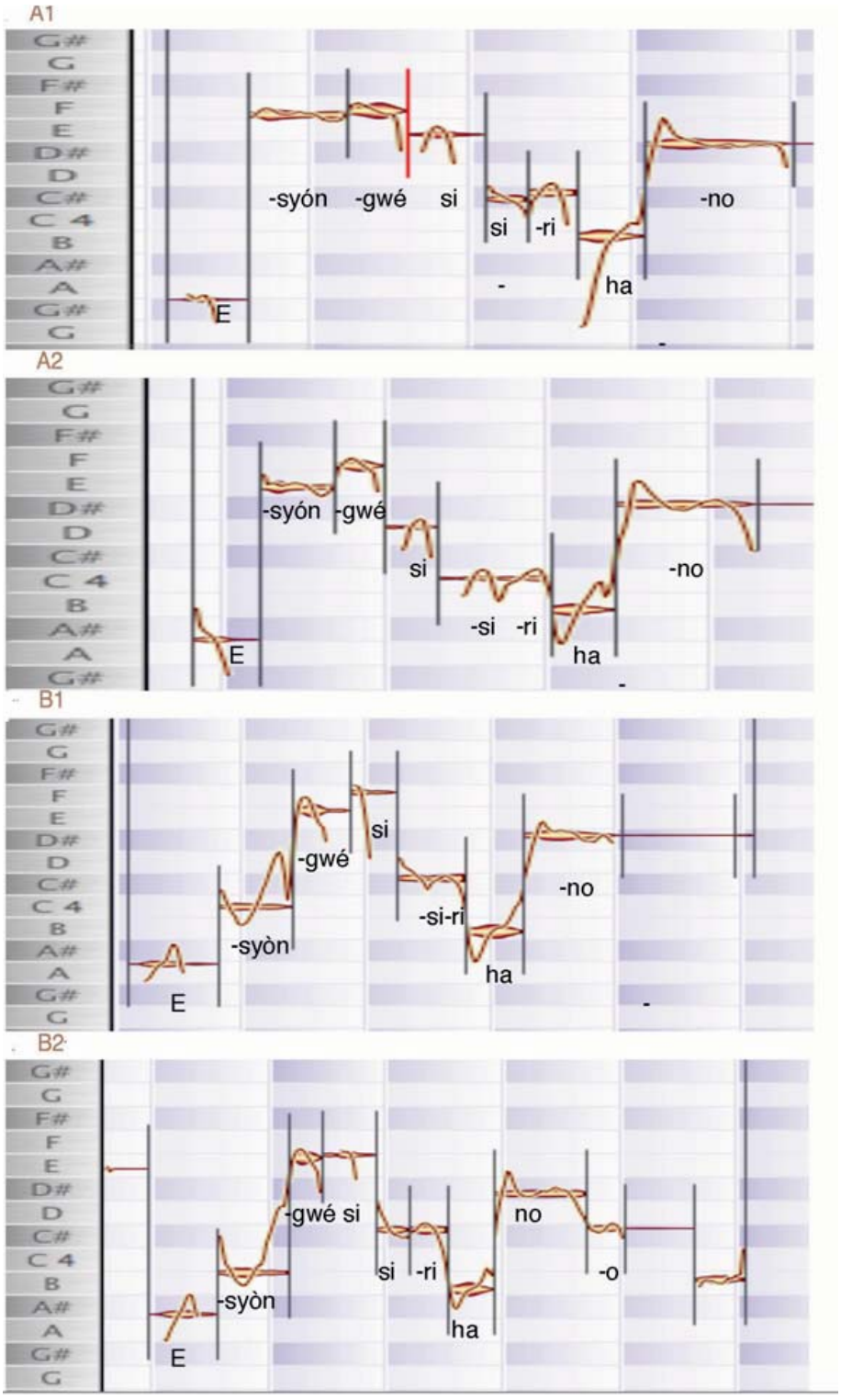


Figure 4: Graphical representation of the sung improvisation performed by musician Erissanya (on the sentence pairs *Esyóngwé: sisiri hano* and *Esyòngwé sisiri hano.* The chart shows the different prosodic profile attributed to the singing at the minimal pair *Esyóngwé* ('wood') and *Esyòngwé* ('leopards'). Software: "Melodyne".

The difference between the two sung sentences is clear in reading the graphs above. The two high tones of the word *esyóngwé* in sentence A are

performed on the same note in the first improvisation (ex. A1) and with a difference of a semitone in the second recurrence (ex. A2), while the word *esyòngwé* of sentence B is performed on an ascending third in version B1 and on a fourth in recurrence B2.

Experiments with other musicians and singers, even on other pairs of sentences, have achieved similar results, even though differences between the tones were expressed with intervals of varying width.

From this experiment we can speculate that each pair of sentences would be sung with different melodies that respected the prosodic profile of each.

Conclusions

The work as a whole has allowed one to trace and highlight influences of spoken language rules on songs. Nevertheless, due to pitch variability of linguistic tones, the melody of a Konzo song appears to retain its autonomy. Indeed, each Konzo sentence/verse/proverb can be sung with different melodies in respect of the prosodic profile. Moreover there are some melodies in the repertoire of Konzo musicians for which new texts are often used. In these circumstances, it is clear that new words must be carefully adapted to the melody of the song, and it is very frequent that associating a text to a melody can require a change in the text or a mutual adjustment of both text and melody. In most cases modern Konzo composers write the lyrics first before they create the melody, or compose lyrics and melody simultaneously. We have developed the hypothesis that the descending profile of sung melodies can be considered in part as derived from the phenomenon of downdrift and that the assonances of the lyrics can be derived from the alliterative concord rule which may affect even the melodic contour of the songs. These aspects, however, were only partially explored in this paper; after all, in order to identify the question of the autonomy of melodies, it has been possible to investigate the topic with only a small sample of data. However, the experimental methodology adopted proves useful for further studies of the relationship between music and tonal languages and provides some possible solutions to many questions still unanswered about this topic. They have occupied the attention of many Africanist ethnomusicologists, and I hope that in the future this challenge will attract more psychologists, linguists and neuroscientists.

Please visit http://www.phonogrammarchiv.at/7DE709JKSAA
for additional appendices.

References

Agamennone, Maurizio & Francesco Giannattasio (ed.). 2002. Sul verso cantato – La poesia orale in una prospettiva etnomusicologica. Padova: Il Poligrafo.

Agawu, Kofi. 1984. "The impact of language on musical composition in Ghana: an introduction of the musical style of Ephraim Amu". Ethnomusicology 28 (1): 37–73.

Agawu, Kofi. 2001. "Musica, lingua e danza nella prospettiva degli Ewe del Nord". In: Nattiez (ed.), Vol. 1, 944–956.

Agawu, Kofi. 2003. Representing African Music: Postcolonial Notes, Queries, Positions. London: Routledge.

Alexandre, Pierre. 1969. "Langages tambourinés: une écriture sonore?" Semiotica 1 (3): 273–281.

Arom, Simha. 1985. Polyphonies et polyrythmies instrumentales d'Afrique Centrale: Structure et Méthodologie. 2 vols. Paris: Société d'Etudes Linguistiques et Anthropologiques de France.

Ayuninjam, Funwi F. 1998. A reference grammar of Mbili. Lanham etc.: University Press of America.

Balinande, Kambale. 2006. Lhukonzo-English dictionary. Kampala: Fountain Publishers.

Banti, Giorgio & Francesco Giannattasio. 1996. "Music and metre in Somali poetry". In: Hayward, Richard J. & Ioan Myrrdin Lewis (eds.). Voice and Power: The Culture of Language in North-East Africa. Essays in Honour of B.W. Andrzejewski. (African Languages and Cultures, suppl. 3). London: SOAS (University of London), 83–127.

Banti, Giorgio & Francesco Giannattasio. 2004. "Poetry". In: Duranti, Alessandro (ed.). A companion to linguistic anthropology. Malden, Mass.: Blackwell, 290–320.

Blacking, John. 1973. How musical is man? Seattle: University of Washington Press.

Cardona, Giorgio R. 1988. Dizionario di Linguistica, Rome: Armando editore.

Carrington, John F. 1949a. Talking drums of Africa. London: Carey Kingsgate.

Carrington, John F. 1949b. A comparative study of some Central African gong-languages. Bruxelles: Falk etc.

Cooke, Peter. 1999. "Was Ssempeke just being kind? Listening to instrumental music in Africa south of the Sahara". The World of Music 41 (1): 73–83.

Cooke, Peter. 2001. "Uganda". In: Sadie, Stanley (ed.). The New Grove Dictionary of Music and Musicians. 2nd ed. Vol. 26. Oxford: University Press, 34–43.

Cooke, Peter & Martin Doornbos. 1982. "Rwenzururu protest songs". Africa 52 (1): 37–60.

Crupi Vanna Viola. 2008. "Roles and function of endara xylophone". In: Pennacini & Wittenberg (eds.), 253–284.

Crupi, Vanna Viola. 2011. Le relazioni fra musica e lingua nella cultura Konzo: repertori vocali, proverbi cantati, generi poetici e narrativi. PhD thesis, Università degli Studi di Roma "La Sapienza".

Crupi, Vanna Viola. 2013. "Music in the sacred forest of the Rwenzori". The Uganda Journal (Special Issue: Sacred Natural Sites and Cultural Heritage in East Africa) 53: 64–89.

Ekwueme, Lazarus. 1974. "Concepts of African Musical Theory". Journal of Black Studies 5 (1): 35–64.

Euba, Akin. 1990. Yorùbá Drumming: The Dùndún Tradition. (Bayreuth African Studies Series, 21/22). Bayreuth: Breitinger.

Facci, Serena. 1988. "La musica dei Wanande dello Zaire". Culture Musicali V/VI (10/11): 87–125.

Facci, Serena. 1996. "I Nande e la loro Musica". In: Remotti, Francesco et al. (eds.). Etnografia Nande III. Torino: Il Segnalibro, 11–57.

Facci, S. & S. Nannyonga-Tamusuza. 2008. "Continuity and Change in Bakonzo Music: From 1906 to 2006". In: Pennacini & Wittenberg (eds.), 223–252.

Gardoncini, S. 2009. "Crinali. Spazio e potere tra i Bakonzo". Africa: Rivista trimestrale di studi e documentazione dell'Istituto italiano per l'Africa e l'Oriente LXIV (3–4): 508–529.

Giannattasio, Francesco (ed.). 1994. Il verso cantato. Atti del Seminario di studi (Aprile–Giugno 1988), Università degli Studi di Roma "La Sapienza".

Giannattasio, Francesco. 2002. "Il tempo delle more: riflessioni sui rapporti fra metrica verbale e ritmo musicale a partire da alcuni esempi di poesia somala". In: Agamennone & Giannattasio (eds.), 137–162.

Giannattasio, Francesco. 2005. "Dal parlato al cantata". In: Nattiez (ed.), Vol. 5, 1003–1036.

Gray, Catherine. 1995. "Compositional Techniques in Roman Catholic Church Music in Uganda". British Journal of Ethnomusicology 4: 135–155.

Guthrie, Malcolm. 1948. The Classification of the Bantu Languages. London: Oxford University Press.

Guthrie, Malcolm. 1971. Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages. Farnborough: Gregg.

Hyman, Larry & Nzama Valinande. 1985. "Globality in the Kinande tone system". In: Goyvearts, Didier (ed.). African Linguistics: Essays in Memory of M.W.K. Semikenke. Amsterdam: John Benjamins, 239–260.

Kavutiraki, Kambale. 1978. Lexique nande-français, français-nande. Kinshasa-Gombe: Editions du Sec. Gén. de l'Episcopat du Zaïre.

Jones, Arthur Morris. 1959. Studies in African music. London: Oxford University Press.

Kubik, Gerhard. 1994. Theory of African Music. Vol. I. Wilhelmshaven: Noetzel.

Kubik, Gerhard. 1999. Africa and the blues. Jackson: University Press of Mississippi.

Kubik, Gerhard. 2010. Theory of African Music. Vol. II. Chicago etc.: University of Chicago Press.

Ladefoged, Peter, Ruth Glick & Clive Criper. 1972. Language in Uganda. London & New York: Oxford University Press.

Mutaka, Ngessimo M. 1994. The Lexical tonology of Kinande. LINCOM Studies in African Linguistics vol. 1.

Mutaka, Ngessimo M. & Kambale Kavutirwaki. 2011. Kinande/Konzo-English dictionary: with an English-Kinande index. Trenton, N.J. etc.: Africa World Press.

Mutaka, Philip. 2008. "The Kinande Bilingual Dictionary: A Tool for Learning about the Nande/Konzo Traditional Culture". In: Pennacini & Wittenberg (eds.), 153–168.

Nannyonga, Sylvia Tamusuza. 2007. "Inculturation of Music: Negotiating the 'Sacred' in the Liturgy of the Roman Catholic Church in Uganda". Paper presented on 15 October 2007 at the Dept. of Musicology, Università degli Studi di Roma "La Sapienza".

Nattiez, Jean-Jacques (ed.). 2001–2005. Enciclopedia della Musica. 5 vols. Torino: Einaudi.

Nespor, Marina. 1994. Le Strutture del Linguaggio. Bologna: Il Mulino.

Nketia, J.H. Kwabena. 1958. "The Contribution of African Culture to Christian Worship". International Review of Mission XLVII (187): 265–278.

Nketia, J.H. Kwabena. 1975. The music of Africa. London: Gollancz.

Patel, Aniruddh. 2008. Music, Language and the Brain. London: Oxford University Press.

Pennacini, Cecilia & Hermann Wittenberg (eds.). 2008. Rwenzori: Histories and Cultures of an African Mountain. Kampala: Fountain Publishers.

Rouget, Gilbert. 1964. "Tons de la langue en gin et tons du tambour". Revue de musicologie 50: 3–29.

Schellenberg, Murray. 2012. "Does Language Determine Music in Tone Languages?" Ethnomusicology 56 (2): 266–278.

Sebeok, Thomas A. & Donna Jean Umiker-Sebeok (eds.). 1976. Speech Surrogates: Drum and Whistle Systems. 2 vols. The Hague & Paris: Mouton.

Stern, Theodore. 1957. "Drum and whistle languages: An analysis of speech surrogates". American Anthropologist 59: 487–506.

Tucker, A. Norman. 1960. "Notes on Konzo". African Language Studies 1: 16–41.

Turchetta, Barbara. 1996. Introduzione alla linguistica antropologica. Milano: Mursia.

Turchetta, Barbara. 2000. La ricerca di campo in linguistica: metodi e tecniche d'indagine. Roma: Carocci.

Turchetta, Barbara. 2008. "Le lingue in Africa Nera". In: Banfi, Emanuele & Nicola Grandi (eds.). Le lingue extraeuropee: Asia e Africa. Roma: Carocci, 489–553.

Turchetta, Barbara (in print). A grammatical sketch of the Lukhonzo language. [MS]

Villepastour, Amanda. 2010. Ancient Text Messages of the Yorùbá Bàtá Drum. Farnham: Ashgate.

Wachsmann, Klaus. 1952. "Musicology in Uganda". The Journal of the Royal Anthropological Institute of Great Britain and Ireland 83 (1): 50–57.

Wachsmann, Klaus. 1953. "The Sound Instruments". In: Trowell, Margaret & Klaus Wachsmann (eds.). Tribal Crafts of Uganda. London etc.: Oxford University Press, 311–414.

Zatorre, Robert. 2003. "Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function". Nature Neuroscience 6 (7): 692–695.

Zemp, Hugo & Christian Kaufmann. 1969. "Pour une transcription automatique des 'langages tambourinés' mélanésiens". L'Homme 9 (2): 38–88.

The realisation of speech tone in Tai Phake music: the case of the Khe Khyang style[1]

(Stephen Morey)

Introduction

This paper will explore two main questions:

> Are the speech tones of Tai Phake realised when singing?
> Is there a concept of melody that is independent of speech and the speech tones among the Tai Phake?

Our study will relate to just one style of Tai Phake song, the *khe khyang* (see below), already discussed in Morey (2010). In answer to (a), we will show further below that at least in this style of Tai Phake song, the speech tones are largely realised.

The answer to (b) is more complex, but as we shall see in towards the end of this paper there is some evidence that an underlying melody, independent of the realisation of the tones, is present in this style.

It is important at the outset to state that in this paper, the term speech tone (or simply tone) will refer to lexical tone; that is tones which distinguish the content meaning of words (see following section for a detailed description of this in Tai Phake). We are not referring to tone over larger linguistic units, such as the phrase or intonation unit (IU). As we will see in Tai Phake, these lexical tones may be a bundle of features including relative pitch, contour (rise and fall), phonation (plain, breathy, creaky and glottal) and duration (short and long).

The Tai Phake

Speakers of Tai Kadai languages have been in the Upper Brahmaputra valley, Assam State, India, for several centuries. The traditional date for the arrival of the first Tai speaking people, the Tai Ahom, was 1228, when a band of Tai led by the prince Siukapha arrived from the Kingdom of Mau

1 This work has been made possible by a grant from the Australian Research Council, under the Future Fellowship scheme, and I am also very grateful to the Centre for Research on Language Diversity, La Trobe University and all my colleagues there. Much help has been gained from many different people in the preparation of this paper; I wish to particularly mention Jürgen Schöpf, Catherine Ingram, Bob Ladd, Amanda Villepastour, Phil Rose, Colleen Holt and Anna Parsons, as well as all the participants of the Workshop on Relationship between Speech Tone and Music, and the anonymous reviewers. My Phake consultants, particularly Aije Let Hailowng and Ee Ngyan Kheit, have both given me a huge amount of help over many years, as well as Am Chaw Khya, Ai Lun Khong, Sam Thon Wingkyen and Ai Chang Chakap.

Lung (now on the border of China and Burma). Following the arrival of the Ahoms, several other Tai groups came to North East India, and settled in villages in Golaghat, Karbi Anglong, Dibrugarh and Tinsukia districts, as well as Lohit District of Arunachal Pradesh. The Tai Phake (*tai² phā⁴ke⁵*) also called simply Phake (Tai Kadai/Southwestern Tai) are a community of around 2000 in 10 villages, and the Phake language is actively used in all of those villages and still learned by children.[2]

The ISO code for Tai Phake is ISO639-3:phk, but Phake is also mutually intelligible with Tai Aiton (ISO639-3:aio) and Tai Khamti (ISO639-3:kht) and to a lesser extent with the highly endangered Tai Khamyang (ISO639-3:ksu). One of the main differences between the four varieties is in the tonal system (see Morey 2005a, b for further details of the other varieties).

Tai languages are notable for their more or less isolating structure, where most words are monosyllabic and consist of a single morpheme. Most Tai languages have a few prefixes that are the conventionalised remnants of full monosyllables, such as Tai Phake /ma¹un⁶/ 'coconut', where /ma¹/ is a reduction of /māk¹/, 'fruit'.

Our work is based on word lists, texts, songs and discussions recorded between 1999 and 2012 in four Phake villages: Namphakey, Borphakey, Tipam and Phanaeng,[3] and on the phonemic and lexicographic work done on Tai Phake by Dr Banchob Bandhumedha in the 1960s and 1970s (Bandhumedha 1987).

Methodological Notes

The issue of how the lexical tone in "tonal languages" is realised in singing has been explored for at least a century. In a seminal paper of 1924, written in English using IPA script, Y. R. Chao pointed out that in at least some styles of Mandarin Chinese music, speech tones were "more suggested than actually represented by the melody composed for the words" and that unaccented syllables "may be set to any note to suit the music". In other words, traditionally at least, tones were realised in song. At the end of the paper he contrasted this with the modern songs "that fit neither the old nor the new system of tones".[4]

2 Superscript numbers for Tai Phake words refer to the tones discussed below.

3 Namphake is located at North 27°16'58", East 95°22'0". Borphakey at N 27°19'41", E 95°42'34". Tipam at N 27°16'45", E 95°23'40. Ninggam at N 27°21'31", E 95°53'30". The recordings discussed here are or will become available through either the DoBeS archive at <http://www.mpi.nl/resources/data/dobes> in the node *Tangsa, Tai, Singpho in North East India*, or through PARADISEC (<http://www.paradisec.org.au/home.html>), or the Phonogrammarchiv of the Austrian Academy of Sciences.

4 I asked my wife, who was brought up speaking several Chinese languages (sometimes referred to as

In our study, we will distinguish three concepts when referring to the relationship of speech tone and music. We will use **realise** to indicate in general terms the way in which speech tone is present in the music, and **represent**, following Chao, to discuss cases where the main features of the speech tone are present. Ladd has suggested three processes whereby tones are **suggested**, to use Chao's term, namely *top-down cues*, which include contextual clues; *phonetic residue*, the use of some associated and possibly less salient features of the phonology of a particular tone to hint at which tone is being used, and *text setting constraints* (see Ladd's contribution in this volume).

Two important papers in the musicological tradition have looked at the realisation of tones in Thai music.[5] Both George List (1961) and Stanley Mendenhall (1975) concluded that there was a correlation between linguistic and musical tones in Thai music, at least in some musical styles. Mendenhall studied a corpus of traditional Thai music collected by William J. Gedney and looking both at the pitch contour of a particular note, as well as the pitches of preceding and following notes, he concluded that the tones were being realised in the melody. To express a falling tone, for example, the pitch may fall during the singing of that syllable, or the following syllable might be on a lower pitch; in the first case the falling tone would be "actually represented" in the melody, and in the second case "suggested", to use Chao's terminology.[6] In a statistical study, Mendenhall found that the contour tones of Thai (rising and falling) were more likely to co-occur with melismas, the rising tone occurring 93.4% of the time with a rising melisma and the falling tone occurring 70.2% of the time with a falling melisma. The three non-contour tones (high, mid, and low), on the other hand, had much high proportions (usually over 50%) occurrence with a non-melismatic structure (called syllabic by Mendenhall).

In this paper we will not be presenting a statistical study, rather we will look in detail at the realisation of tones in citation and show that at least for some styles of Tai Phake, the tones are realised when singing and that this can be demonstrated by instrumental studies.

'dialects') in Malaysia about this. When she was small she remembers being confused about the meaning when listening to modern Mandarin Chinese songs, because, as she told me "the tones changed", that is words did not appear to have the same tone in song as they did in speech.

5 The term Tai is used for the whole language family and for the varieties spoken in India, Burma and SW China. The term Thai is used for the language of Thailand. The two words are cognate.

6 I have often heard it said that (popular?) songs in Thailand do not realise the tones. After a careful study, Ware (2006 and pers. comm.) concluded that in popular Thai music tones are realised.

Tai Phake Tones

The study of tone in Tai languages rests on a fairly secure historical and comparative basis. All of the spoken Tai languages are tonal, having generally between three and seven tones for words in citation, based on tonal categories that can be related back to proto-Tai.

In our discussion of Tai tones, we will use *pitch* to refer to a relative pitch, and *contour* to changes in that pitch within the tone. Absolute pitch is not a useful category for tones within language, because obviously the high pitch range of one individual is not the same as that for another. Most men, for example, have lower absolute pitch than most women. However, we do want to draw a distinction between pitch (high, mid, low) and the contours (rising, falling, rising then falling and level). Most particularly we want to argue that level is a contour feature.

Li (1977) and Gedney (1972) posited that Tai languages had three proto-tone categories for open syllables (vowel or nasal final), conventionally called A, B and C, and two categories for closed syllables (stop final) conventionally called D1 and D2 (long and short respectively). The proto-tones are thought to have been distinguished by phonation (plain, breathiness and creakiness) rather than by pitch or contour, and in particular the C tone is reconstructed as having creaky phonation. In their modern reflexes, these tones are often categorised by pitch and contour distinctions (see Morey 2005b for a discussion of the historical tonology of the Tai languages in North East India).

The modern reflexes of these tones come about as a result of tone changes that were influenced by a combination of the proto tones and the proto initials in four series: voiceless aspirated, voiceless unaspirated, pre-glottalised and voiced.[7] As an example of how this is applied, in most Tai languages the originally voiced stops and fricatives have become voiceless, usually voiceless unaspirated in Shan and the Tai languages of North East India (e.g. *d > /t/), while becoming voiceless aspirated in standard Bangkok Thai (e.g. *d > /th/). An example of this change is the word Tai itself, realised as Thai with aspirate /th/ in Thailand. Furthermore in many Tai languages the stop-fricative distinction has also been lost, and proto fricatives are realised as stops. This would create ambiguities if not for the so-called "tone splits". Consider the Phake words /khā^{2}/ 'thatch' and /khā^{6}/ 'leg'. In proto-Tai these are reconstructed by Li (1977) as *ɣa and

7 The first three are often grouped together as the 'voiceless series' and the last as 'voiced series'. Thus proto words are often given tone indications as A1 or A2, where A refers to the proto tone and 1 to the voiceless series and 2 to the voiced series. In Southwestern Tai, however, the three series of voiceless initials have to be posited to explain the variety of tonal systems in the modern languages.

**kha* respectively, where the distinction is one of initial consonant.[8] Any subsequent merger of those consonants could have created homophones, but the tone split meant that the group of words that had an originally voiced initial with tone A have become Tone 2 (high rising then falling) and the group of words that had an originally voiceless initial with tone A have become Tone 6 (rising).

Bandhumedha (1987) analysed Tai Phake as having six tones, and gave the following description and exemplification of the tones. We will term these citation tones, because, as we will see later, the tone can sometimes alter.

1	even tone with a slight rising at the end	mā¹	'shoulder'
2	high tone	mā²	'to come'
3	grave tone, with a glottal stop	mā³	'mad'
4	falling tone	mit⁴	'seeds, knife'
5	grave tone	mā⁵	'not'
6	high rising tone	mā⁶	'dog'

Table 1: Phake citation tones following Bandhumedha (1987)

Our work suggests that while the categories proposed by Bandhumedha are robust, the form of the tones may have changed over time. For example, we have found that Tone 3 is best described as creaky throughout the syllable, rather than having a final glottal stop, although the latter phenomenon is found when this tone is exemplified in Ngi Pe Pang's *Khe Khyang* song (example 18 below). Also, we treat Tone 2 as a high falling tone, even rising and falling, because the fall is the most prominent feature of the present day realisation of the tone. Similarly in the speech of some modern consultants, Tone 5 is a long low falling tone.

Table 2 presents the tones as spoken by Ee Kya Gohain, a female speaker aged around 35 at the time of recording. In the analysis we present here, the four tones combine the features of (1) pitch (high, low), (2) contour (level, rising, falling), (3) duration (short, long) and (4) phonation (plain, creaky). The key features for each tone are listed in the second column of Table 2.

8 Li's reconstructions along with several others, can be searched at <http://sealang.net/crcl/proto/>.

No.	Tonal features	Examples	Assamese
1	level	nā¹ ‘quarrel’	kaziya
2	high rising then falling	nā² ‘rice field’	pothar
3	creaky	nā³ ‘face’	mukh
4	falling	nā⁴ ‘mother’s younger sister’	mahi
5	low (falling) and long	nā⁵ ‘melt away’	goli juwa
6	rising	nā⁶ ‘thick’	dath

Table 2: Tones 1–6 pronounced on the syllable *na* in Tai Phake[9]
SDM01_2007_019_NaTones.wav

An intensity and pitch trace[10] for each of these tones given in citation (using the words in Table 2) is given in Table 3.

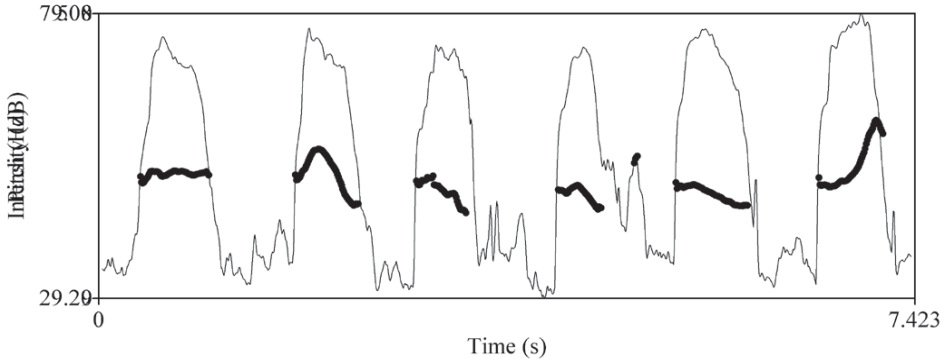


Table 3: Pitch trace for the 6 Phake tones exemplified in Table 2

As can be seen in this table, the Tones 1, 2, 5 and 6 are longer than Tones 3 and 4. This is a reflex of the proto-tone, because Tones 3 and 4 in Phake are reflexes of the C tone, generally believed to have been originally creaky or glottalised (see Morey 2005b) and this explains their shortness. Creakiness is still very much in evidence with Tone 3 as realised by Phake speakers today, whereas its reflex in Tone 4 lacks the creakiness and the most salient inheritance of the proto tone is its shorter duration, but there is also some final breathiness in the recording on which Table 2 is based. In other Tai varieties this pattern is reversed, in our analysis of Khamti, for example, creakiness is present in the cognate of Phake Tone 4, whereas the cognate of Phake Tone 3 merely has final glottal constriction (see Morey 2005b).

9 The phonetic symbols used here vary slightly from IPA standard. In Tai Phake, following Bandhumedha (1987), macrons are used to show vowel length, as /ā/. Aspiration is shown by following <h> as /th/, /ph/ and the palatal semivowel is shown as /y/. We also use Bandhumedha’s system of marking tones.

10 All of our instrumental studies are done using PRAAT (<http://www.fon.hum.uva.nl/praat/download_win.html>). In this paper all pitch measurements from PRAAT are given in two forms, Hz values (rounded to the nearest whole number) and semitone values where 440 Hz is set as 0 (rounded to two decimal points, in other words ‘cents’).

The other four tones (1, 2, 5 and 6) are all longer in duration, as is clear from both the intensity and pitch traces. All four of these tones also have plain phonation,[11] so in the case of these tones the distinguishing features are pitch and contour alone. For these tones the use of 'pitch numbers', where 1 is low, 2 mid-low, 3 mid, 4 mid-high and 5 high, could be employed, as shown in (1), though in the form given here this might suggest that some of these tones are shorter in duration than others and that is not the case.

1) Tone 1 44
   Tone 2 452
   Tone 5 31
   Tone 6 335

This notation does not work for distinction of Tones 3, 4 and 5, because they have almost identical pitch traces and would have to be notated as 31, and we would have the following:

2) Tone 3 31
   Tone 4 31
   Tone 5 31

This is demonstrated clearly when we compare just the pitch and contour of the tones, using Hz and semi-tone values, based on the recording exemplified in Table 4:

No.	(1) Pitch (Hz)	Pitch (Semitones)	(2) contour
Tone 1	218-219	-12.10 -12.04	level
Tone 2	233-260-165	-10.98 -9.10 -16.90	high then falling[12]
Tone 3	198-155	-13.81 -18.04	falling
Tone 4	197-159	-13.86 -17.57	falling
Tone 5	198-163	-13.81 -17.16	falling
Tone 6	195-219-309	-14.04 -12.07 -6.11	rising

Table 4: Tones 1-6: Pitch and contour

On the basis of Table 4, we can see that with only features (1) pitch and (2) contour, tones 3-5 could not be contrasted. The feature of (3) duration can also be measured, as in Table 5, and this demonstrates that while there is no pitch/contour difference between tones 4 and 5, there is a duration difference.

11 One of the reviewers of this paper suggested the term 'modal' rather than 'plain' for the unmarked phonation. We have retained the term 'plain' because it conveys more clearly that this is the unmarked phonation type, used in five out of the six tones.

12 Given that the initial consonant of this tone is a voiced nasal [n], it would be expected that the pitch would rise, however at least some of this rise may be associated with the tone, as it continues to rise in pitch after the commencement of the vowel.

No.	Duration (seconds)
Tone 1	0.63”
Tone 2	0.57”
Tone 3	0.44”
Tone 4	0.36”
Tone 5	0.66”
Tone 6	0.56”

Table 5: Tones 1-6: Duration

Historically the difference between tones 3 and 4 on one hand, and the others, was one of phonation. In the modern language, however, the creaky phonation that we suggest was historically present in tone 4 is no longer present, and remains only in the form of shorter duration. Thus a new contrast is created, one of duration.

The final contrast, between tones 3 and 4, is that of phonation. Tone 3 is creaky and tone 4 is plain.

The interplay of the four types of contrast is summarised in Table 6, where the features we take to be contrastive are shown in bold.

	(1) Pitch	(2) Contour	(3) Duration	(4) Phonation
Tone 1	Mid-High 218-219 Hz	**level**	long	plain
Tone 2	Mid-**High**-Low 233-260-165 Hz	**high then falling**	long	plain
Tone 3	Mid-Low 198-155 Hz	falling	short	**creaky**
Tone 4	Mid-Low 197-159 Hz	**falling**	**short**	plain
Tone 5	Mid-**Low** 198-163	falling	**long**	plain
Tone 6	Mid-High 195-219-309 Hz	**rising**	long	plain

Table 6: Phake Tones 1-6: Summary

From this, we can say that pitch is arguably not a key feature of any of these tones, although we have bolded pitch features for Tone 2 and Tone 5. What is clear is that relative pitch alone is certainly not **the sole** key feature for any of them. For this reason, rather than using the pitch number notations seen in (1) above, following Banchob (1987) we are using superscripted numbers to indicate the tone categories that are defined in Table 2 above.

There is a further complicating factor. Phake as analysed by Bandhumedha (1987) and Morey (2005a: 126) has length contrast between /a/ and /aː/ (long /aː/ being written as <ā>), most frequently in syllables with

final consonants: nasals, semi-vowels and stops.[13] Therefore duration can be treated as a feature either of the vowel and/or of the tone.

At first glance this might suggest a confusion between words having Tone 4 with a long vowel and Tone 5 with a short vowel, two such words being /kān^4/ 'break off' and /kan^5/ 'directions'. In each case the coda /an/ is approximately the same length, however the difference is that where there is a long vowel and short tone, the vowel /a/ takes up a higher proportion of the coda, whereas with a short vowel and longer tone, the final nasal /n/ takes up a higher proportion of the coda. This can be seen in both the intensity traces and formant diagrams for these two words in Table 7.

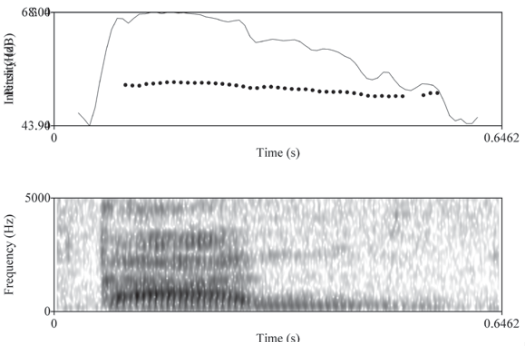
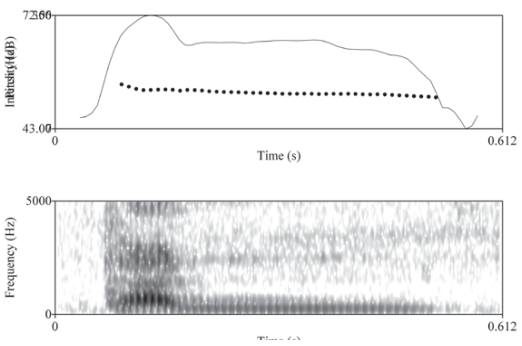
/kān^4/ 'break off'	/kan^5/ 'directions'
	

Table 7: Duration in tone and vowel
phk_20120302_01_SM_H4n_AijeLet_VowelsAndTones_KanTones.wav
phk_20120302_01_SM_H4n_AijeLet_VowelsAndTones_KaanTones_edited.wav

This demonstrates a crucial observation about tone in Tai Phake, namely that it is a feature of the syllable, not of the vowel. In other words the syllable is the tone bearing unit (or perhaps the coda) rather than the vowel. This is one of the reasons why we do not use the IPA system of tone marking, using diacritics on the vowel, because such a system could be read as suggesting that tone is a feature of the vowel alone.

With vowels other than /a/, the duration contrast between words bearing Tone 4 and those bearing Tone 5 is solely an issue of tone, as with the distinction between /yɔ^4/ 'praise' and /yɔ^5/ 'loosen'. This is also the case with words having final /-a/, because there is no vowel length contrast in this situation.

We have mentioned above that proto-Tai is reconstructed as having three tones on open syllables and two (one short and one long) on closed, or stop

13 In vowel final syllables (words) that have /a/, it is always realised long – with one exception. The reduced syllables that are used in conventionalised compounds, such as *ma^{1}un^6* 'coconut' is short and is notated with Tone 1, though the realisation of the tone is a little different.

final syllables. As in the proto-language, there are fewer tonal categories with closed syllables than with open syllables. Closed syllables in citation in Phake are found with Tone 1 and Tone 4 only, and there is no length distinction on closed syllables in Phake.[14]

These tonal categories in Phake are well recognised by Phake speakers. Following contact with Dr Banchob Bandhumedha in the 1960s and 1970s, and more recently with Stephen Morey (since 1996), many of the Phake elders and others interested in language issues are accustomed to referring to the tones using the numbering system in Table 2. For example, when asked to categorise the tone of an unfamiliar word, some Phake speakers will spell out all of the possibilities in the order Tone 1 through to Tone 6. Consider the following exchange between Stephen Morey and Aije Let Hailowng, a male speaker aged around 75 at the time of recording. In this example Aije Let is asked about the tone of the negative form of the verb 'to eat'.

SM

3)	ma^{5}	kin^{6}	caü^{2}	nā^{1}	
	NEG	NEG.eat	yes	QN[15]	

'"Not eat", isn't it'

Aije Let

4)	ə^{2}	ma^{5}	kin^{6}	caü^{2}	ū^{1}
	EXCL	NEG	NEG.eat	yes	STAY

'Yes, it is "Not eat".'

SM

5)	ma^{5}	kin^{6}	seŋ^{6}	saŋ^{6}	pen^{2}
	NEG	NEG.eat	sound	which	be

'And "Not eat" is which tone?'

14 There are a small number of words with closed syllables that Bandhumedha (1987) notates with the 2nd or 3rd tone. Most of these are onomatopoeic and are not reconstructable to proto-Tai. Words with closed syllables that are reflexes of proto-Tai are found with only the 1st and 4th tones.

15 An accidental use of the Aiton form of the question particle, in place of Phake /nɔ^{6}/.

Aije Let

6) kin⁶ / kin¹ kin² kin³ kin⁴ kin⁵ kin⁶ kin⁶ / hok¹

NEG.eat NEG.eat six

‘“Not eat?” *kin¹ kin² kin³ kin⁴ kin⁵ kin⁶*, it is number six.’

AiCheLet_ToneChanges_Negative.wav

Another important feature of Tai Phake, also demonstrated in (3) to (6), is the presence of what we will term changed tones. For example, some verbs, in most cases those that carry the Tone 2 (high rising then falling), such as /kin²/ ‘be’ are realised with a rising tone, the same shape as Tone 6, when they are negated. This is not found in other Tai languages like Tai Aiton. Tone change processes in Phake are used for negation, imperatives and questions, and are described in detail in Morey (2005b).

Khe Khyang style

There are a number of distinct styles of song in Tai Phake (see Morey 2010 for a list of some of these), all of which are grouped together as /mɔ⁶ khām²/ (literally ‘sing word’). In this paper we will be closely examining just one of these, the /mɔ⁶ khām² khe² khyāŋ²/ (မော်ၵံမေၶျင်), a song usually written to mark some kind of event. We will examine three versions of the song:

7) i) *Khe Khyang* written and sung by Aije Let Hailowng (born c. 1935), to mark the publication of the Tai primer produced by Stephen Morey, performed in January 2000[16]

ii) *Khe Khyang* written and sung by the late Ngi Pe Pang (c. 1920–1980) to mark the visit of Thai scholar Banchob Bandhumedha, and recorded by her in about 1970[17]

iii) *Khe Khyang* written and sung by Am Chaw Khya (born c. 1960), to mark the publication of the Tai primer produced by Stephen Morey, performed in January 2000[18]

16 The original recording was made in noisy circumstances and so in March 2001 a second recording was made on Minidisc. This recording was made in mono, and is archived at PARADISEC under the name SDM01-20010102-001.wav. A discussion of the meaning of the text was recorded under the name SDM01-20010102-002.wav.

17 The original cassettes of these recordings were kindly loaned to us by Dr Navavan Bandhumedha, niece of the late Dr Banchob Bandhumedha. They are currently being digitised and will be archived at both the Phonogrammarchiv of the Austrian Academy of Sciences and also at PARADISEC. The digitised files have yet to be named.

18 Like the *Khe Khyang* song of Aije Let, the original recording was made in noisy circumstances and so in November 2003 a new version was recorded by video. The digitised file of this video was named as SDM01-20000129-03_NamphakeyMeeting3_AmChawSong. It was re-recorded under better conditions as SDM01-20031111_09_AmChawKhya_KhamKheKhyang.mpeg. Both recordings will be archived at PARADISEC.

The *Khe Khyang* style can be described as a kind of a *Sprechstimme*. Rhythmically it is arranged in lines of seven syllables, which have a strict rhyming scheme that links each line to the next.

The 4th and 7th syllables/notes of each line are certainly the longest and most stressed of the syllables and to me, impressionistically at least, the other syllables/notes are like upbeats to them.

The rhythm seems to follow the durational pattern shown in (8), which is represented by the wave form in Table 8, an example of Aije Let singing the tune of a line of the *Khe Khyang* song without words, on the vocable syllable *ta*.

8) short half-long short long; short half-long long

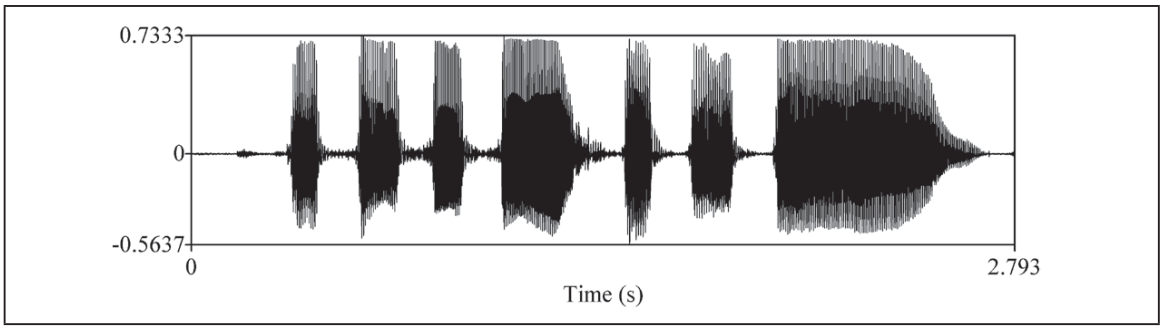


Table 8: Rhythm of *Khe Khyang* style
Khe_Khyang_discussion_Line33_WithoutWords.wav

The longer duration of the 4th and 7th or last syllabels can clearly be seen in table 8. The fact that the first syllables of each section of the line, the 1st and 5th syllables, are shorter can also be seen. Since duration is a feature of the tonal system, as already mentioned above, the presence of an underlying duration feature in the rhythmic structure will have consequences for the realisation of tones.

The rhyming system of the *Khe Khyang* style is shown in (9), with superscripts to mark the tones required as part of the rhyming system. This type of rhyme is called 'waist rhymes' where the last syllable of a given line (termed /ta^{1} phū^{3}/ 'male' in Tai Phake) rhymes with the fourth syllable of the next line (termed /ta^{1} mɛ^{5}/ 'mother'). These two syllables will be termed rhyming positions. I have used capital letters to show that the last syllable of the first line (marked 'A') rhymes with the fourth syllable of the second line (also 'A') and that the last syllable of the second line rhymes (marked 'B') rhymes with the fourth in line 3 (also 'B') and so on.

9) ♪ ♪ ♪ ♪ - ♪ ♪ A♪ ^{1}
♪ ♪ ♪ A♪ ^{1} - ♪ ♪ B♪ ^{2}
♪ ♪ ♪ B♪ ^{2} - ♪ ♪ C♪ ^{1}
♪ ♪ ♪ C♪ ^{1} - ♪ ♪ D♪ ^{2}

In Aije Let's *Khe Khyang*, the tonal pattern of a /ta^{1} phū^{3}/ on Tone 1 alternating with one on Tone 2 at the end of each line continues throughout the whole song. This pattern is not followed for all *Khe Khyang* songs, as we will see below. The important point is that tone (with its concomitant pitch and contour) on the 4th and 7th syllables is part of the rhyming system, and arguably not part of any underlying melody, at least for the words in rhyming positions.

Below we will discuss in more detail the evidence for an 'underlying' melody in this style. We have recorded the song without words, sung on a vocable *ta*. When this was done, the melody of the pair of lines was as in (10), which employs the cipher notation, where 1 represents the 'tonal centre' of this melody, which is also the pitch of the level 1st tone in the rhyming position.

10)	♪	♪	♪	♪	-	♪	♪	♪
	1	1	2	2 F		6̣	5̣#	1
	♪	♪	♪	♪	-	♪	♪	♪
	2	2	6̣	1		7̣	1	2 F

In (10), the notation 2 F represents a note that is sung on a falling tone (Tone 2). Since a fall is required by the rhyming system at this point, we are unable to distinguish any underlying melody from the realisation of the tone.

The realisation of speech tones in the *Khe Khyang* style

As already explained, in Aije Let's version of *Khe Khyang*, the speech tone and musical pitch of the 4th and 7th syllables of each line is fixed. When the 7th syllable of a line is a word with Tone 1, the melody has a level pitch at the tonal centre; the syllable that rhymes with this will similarly have a melody with level pitch at the tonal centre; this is given in cipher notation as 1. Where the 7th syllable of a line is a word with Tone 2, the melody commences on a level pitch around two semitones above the tonal centre and then falls; this is given in cipher notation as 2 F. These two positions in each line are places where the tone is fully represented in the musical melody, and the melody is fixed at that point.

We will now investigate whether tones in other positions within the line are also represented, and we will commence our examination by looking at the occurrence of Tone 6 (rising) when it occurs in the 6th (second last) syllable. In examples (11), (12), (13) and (14), the same word, /lāi^{6}/ 'many' is found in the 6th syllable. Since this is not one of the positions that carries a fixed tone, we will first investigate two questions:[19]

19 Detailed pitch analyses of each of the examples discussed in Section 4 are presented in the Appendix.

Does the rising tone (Tone 6) get realised in this position, and What happens to the realisation of the 5th (third last) syllable, which differs in form and tone, across the four examples?

11) တီ ထၢၼ် ၼဲ ၼၢၼ် လိၵ် လဵဝ် ထၢၼ်

tī^{5}	than^{4}	nai^{3}	ñān^{1}	lik^{4}	lāi^{6}	phan^{2}
at	there	get	knowledge	book	many	type

'There he acquired much and various knowledge.'

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (10)
SDM01-20010102-001_10.wav

12) ၽူ ၶၢင် ၸေ လိတ် လိၵ် လဵဝ် ထၢၼ်

hū^{4}	cāŋ^{5}	phɛ^{5}	phɛt^{1}	lik^{4}	lāi^{6}	phan^{2}
know	know	increase	scatter	book	many	type

'He knows many things, of many types.'

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (12)
SDM01-20010102-001_12.wav

Table 9 gives the pitch and intensity traces for examples (11) and (12) with the rising tone on the word /lāi^{6}/ 'many' marked by an arrow:

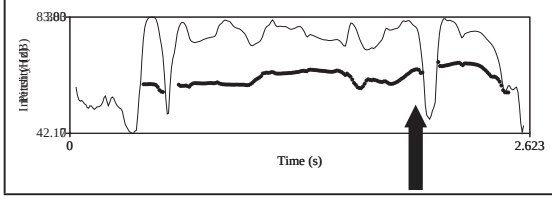
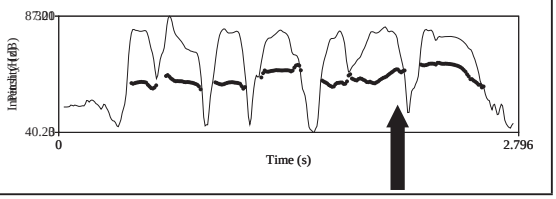
Example (11)	Example (12)
	

Table 9: Demonstrating the realisation of tones in the *Khe Khyang* style (1)

13) ၼၤ် မိင် ၼဲ ၵူ စဲ ဃံ ထၢၼ်

naü^{2}	məŋ^{2}	nai^{4}	ū^{1}	tai^{2}	sām^{6}	phan^{2}
in	country	this	stay	Tai	three	type

'There are three tribes of Tai in this country.'

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (84)
SDM01-20010102-001_84.wav

14) တီ ၶဝ် ထံ ၶဝ် ၶၢ လဵဝ် ပုင်

tī^{5}	khau^{6}	thām^{6}	au^{2}	khɔ^{3}	lāī^{6}	puŋ^{1}
at	3PL	ask	take	subject	many	story/meaning

'"Ask the (Tais) and take much information and many stories.'

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (33)
SDM01-20010102-001_33.wav

Table 10 gives intensity and pitch traces for examples (13) and (14):

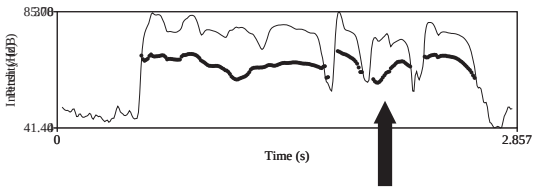
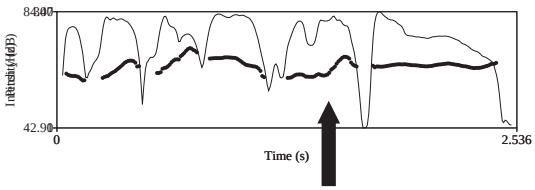
Example (13)	Example (14)
	

Table 10: Demonstrating the realisation of tones in the *Khe Khyang* style (2)

Table 9 and Table 10 show the pitch trace for the word /lāi^{6}/ in each of these examples, and as we can see, there is a rise in every case. Notice that in (11), (12) and (13), which are examples from the second line of the pair of lines, the last syllable shows the pronounced fall associated with Tone 2, whereas (14) is an example from the first of the pair, and here the last syllable shows the long and level pitch associated with the Tone 1. Nevertheless, in each case rise is present. This rise, we suggest, is the realisation of tone and is not a feature of any (underlying) musical melody.

To show that this rise is not a necessary part of the melody, but rather is an example of tone being marked within the restrictions of the melodic structure of this song, we will compare with (15), an example where the 6th or second last syllable does not have a rising tone, but a falling tone as we can see from the intensity and pitch trace given in Table 11, showing a clear fall in pitch across this syllable

15)

ယ့	တပ်	ဒိၚ်	စဳ	သွံ	ခှံ	ကော့
yā^{1}	tap^{4}	ŋin^{2}	wāŋ^{1}	thɔm^{1}	khām^{2}	kau^{2}
don't	must not	listen	confused	listen	word	1SG

"'Don't be confused, listen to my words'".

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (32)
SDM01-20010102-001_32.wav

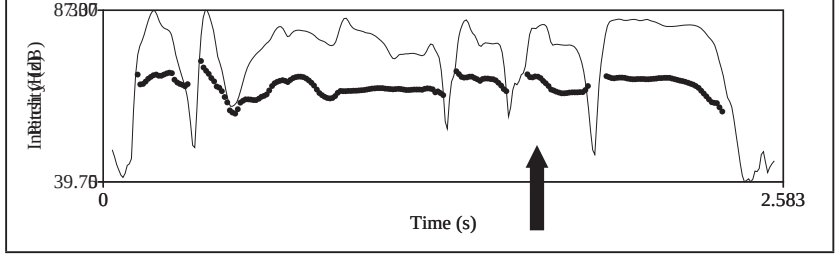
Example (15)


Table 11: Demonstrating the realisation of tones in the *Khe Khyang* style (3)

Returning to Table 9, we notice that in examples (11) and (12), the 5th or third last syllable is a short consonant final word carrying the Tone 4 (falling). The fall is not particularly clear in the pitch trace, but is clearly distinguishable from (13), where the 5th or third last syllable is carrying the Tone 2 (high falling).

Example (14) is more complex because the 5th or third last syllable is the word /khɔ³/ 'subject', which carries the Tone 3 (creaky). In citation this tone is usually falling, but the key feature, as we argued above, is creakiness. As we will see below in (17) and (18), when Tone 3 is in the rhyming position, it is realised with a final glottal constriction, but that is not apparent in (14).

Impressionistically we do hear a very slight amount of creakiness in the voice of the singer on this word. The 5th syllable is demonstrably short, when compared for instance with the 5th syllables in examples (13) and (15).

Perhaps shortness is the only cue here; but that would create a potential ambiguity with a word of the form /khɔ⁴/, possibly meaning 'irritating, be diseased', a far less frequently encountered word than /khɔ³/ 'subject', and one that would not make sense in this situation.

It thus appears to be the case that Tai Phakes also rely on context to disambiguate the meaning, and that consequently tone does not always have to be explicitly realised. The only sensible meaning of /khɔ/ in line (14) is 'subject', and the context of the surrounding words will make this clear. Aije Let specifically said that this was how the meaning of the word can be disambiguated in some cases. In January 2010, we had interviewed him about how a certain word appeared to have a changed tone in song; the phrase /khun⁶ e⁶ nāŋ²/ ('prince' 'wish' 'lady'), translated as 'I, the prince wish for the lady', was realised in a song sung by E Ngyan Kheit with a clear level tone on *e*, in other words Tone 1, in place of Tone 6. In 2010 he had suggested that this should be regarded as an example of tone change, but when I visited him again in March 2012, he was keen to correct what he thought was an error, saying of this very example that: "As for the difference between *e*⁶ and *e*¹, which you asked about [two years before] ... how do we understand it? Taking (the meaning) of the words behind and in front, then we will get to understand."

We would suggest that this methodology relies on at least some of the words being clear at first hearing; this would be true of /au²/ and /lāī⁶ puŋ¹/ in (14), and from this context, where the tones are clearly realised, the meaning of /khɔ³/ can be deduced. Let us now examine again (15), whose pitch trace is in Table 11. Each of the syllables is realised with pitch traces that match the expected pitch movement for the tones, as shown in (16):

16)

1st syllable	Tone 1	level
2nd syllable	Tone 4	direct fall, short (stop final)
3rd syllable	Tone 2	rising then falling
4th syllable	Tone 1	level
5th syllable	Tone 1	level (a little fall at the end)
6th syllable	Tone 2	high falling
7th syllable	Tone 2	level then falling

The presence of all of these contours is not found in the example given in Table 12, where Aije Let was singing an example of the 1st line of a pair but without meaningful words, on a vocable *ta*. As can be seen here, each of the syllables is sung with a level tone, except for syllable 4 (which carries Tone 2, high falling) and also syllable 6, the second last syllable. The rise in this latter syllable may be explained either by the need to reach a higher pitch, or the fact that this line was sung after a discussion of line (14) where the second last syllable is /lāi^{6} / 'many', pronounced on a rising tone.

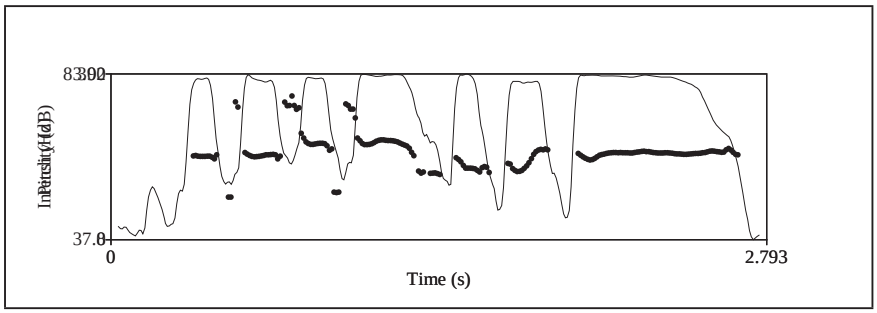


Table 12: A line of *Khe Khyang* sung without words
Khe_Khyang_discussion_Line33_WithoutWords.wav

Having established that at least some of the tones are fully represented in the song, let us examine another *Khe Khyang* in which Tones other than Tones 1 and 2 can be used in rhyming position. An older, but somewhat shorter *Khe Khyang* was composed by the late Ngi Pe Pang[20] in honour of Dr Bandhumedha in the 1970s. Unlike Aije Let's *Khe Khyang*, the pattern of tone rhyme varies a little from the regular Tone 1 ~ Tone 2 alternation, as we can see in Table 13:

Line Nos.	1st line	2nd line
2-3	Tone 4	Tone 1
4-5	Tone 4	Tone 1

20 When I got copies of these recordings, I took them to Namphakey village in March 2001. As I started playing the recording, E Ngyan Kheit spoke to her son, Sam, and he got up and walked out – I assumed because he did not like listening to traditional songs. However I was wrong. He soon returned with the late Ngi Pe Pang's son who had been aged 12 months when his father died and thus had, for the first time, the opportunity to hear his father's voice.

6-7	Tone 2	Tone 4
8-9	Tone 1	Tone 2
10-11	Tone 1	Tone 2
12-13	Tone 1	Tone 3
14-15	Tone 1	Tone 2
16-17	Tone 1	Tone 2
18-19	Tone 1	Tone 2
20-21	Tone 1	Tone 2

Table 13: Tone rhymes in Ngi Pe Pang's *Khe Khyang*

Of particular interest is the use of Tone 3, as in (17) and (18). In (17), the distinctive creaky Tone 3 is found in the final rhyme leading position (Phake /ta^{1} phū^{3}/), a tone repeated in line 14, whose translation is presented as (18):

17)

ဂျိပ်တိုက်	ဃုပ်	လေ	စွက်	မွက်ယျ
khɛp^{1}tin^{2}	sup^{1}	lɛ^{1}	tɔk^{1}	mɔk^{1}yā^{3}
shoe	wear	roam	press	flower

Roaming, wearing shoes with a flower pattern.'

18)

မ	ဂူ	တီ	ဂ	ဂုပ်	ဃီကွက်
mā^{2}	nū^{2}	tī^{5}	nā^{3}	khup^{1}	sī^{4}kɔn^{1}
come	look	place	face	very nice	beautiful

'Come to look at her beautiful face.'

Song in the Khe Khyang Style – to honour Dr Banchob Bandhumedha,
sung by the late Ngi Pe Pang, Nos. (13) and (14)
Banchob_2_5_NgiPePang_KhamKheKhyang_13and14.wav

The creakiness feature of this tone is audible in the song as a final glottal, not apparent in the lines preceding and following it, where the final element carries Tone 1 (Level tone). We have included a full wave form here to show the abrupt ending of the last syllable, quite distinct from the tapering ending that we see in the two examples from Am Chaw Khya in Table 15 below:

example (17)

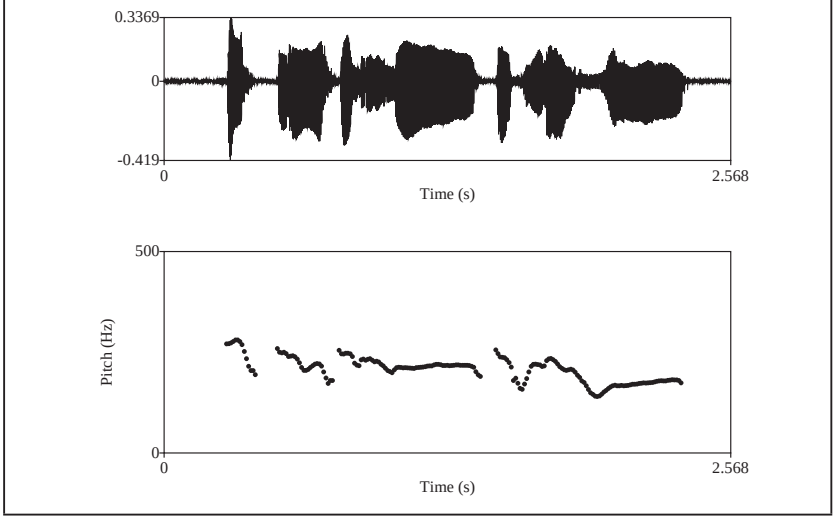


Table 14: Demonstrating the realisation of tones in the *Khe Khyang* style (4)

We conclude from this that Tone 3 is actually realised here, by the presence of the final glottal, and sounds quite different in rhyming position from the other tones. The fact that the realisation of the tone is indicated by the final glottal rather than a fully creaky tone, is an example of what Ladd (this volume) calls phonetic residue.

In the *Khe Khyang* songs we have investigated so far, only Tones 1 to 3 have been used in rhyming position, but in Am Chaw Khya's song Tone 5 is employed, in the following pair:

19)

ꪀꪱ	ꪜꪶꪀ	ꪜꪊ	ꪮ	ꪚꪱ	ꪻꪀꪎ
haü^{3}	lik^{4}	pɔ^{2}	mī^{2}	tai^{2}	phā^{4} ke^{5}
give	book	enough	be	Tai	Phake

'As this book is for the Tai Phake.'

20)

ꪀꪱ	ꪉꪉ	ꪁꪱ	ꪎ	ꪫꪱꪀ	ꪀꪊ
haü^{3}	hup^{4}	caü^{2}	ŋɛ^{5}	nüŋ^{1} nai^{4}	kā^{1}
give	collect	mind	mild	like this	GO

'Let us be broad-minded like this.'

Song in the Khe Khyang Style – to honour Stephen Morey, sung by Am Chaw Khya, Nos. (39) and (40)
PhakeSongs_AmSaeuKheKhyang_39and40.wav

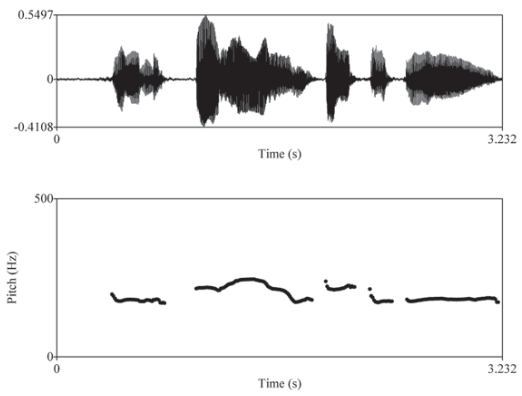
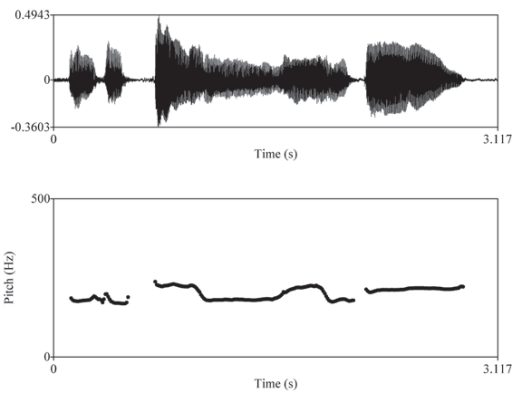
Example (19)	Example (20)
	

Table 15: Demonstrating the realisation of tones in the *Khe Khyang* style (5)

The difference between Tone 5 and Tone 1 is conveyed by pitch, with the words carrying Tone 5 (marked by arrows in Table 15) being of a lower pitch. Tone 6 has not been recorded in a rhyming position.

Finally, let us investigate whether the duration contrast between Tone 4 and Tone 5, that we see in E Kya's pronunciation of the citation tones, can be demonstrated in the *Khe Khyang* style. To do this, we want to find words in the same position with the syllable structure carrying the same vowel, but differing only in terms of citation tone. Consider (21) and (22), for which the intensity and pitch traces are given in Table 16.

21)

ပွာ်	ထိက်	ၼဴ	မိုင်	တီ	ထဲလိၵ်ာ်
sɔn^{6}	lik^{4}	naü^{2}	məŋ^{2}	tī^{5}	thai^{1} lɛn^{1}
teach	book	in	country	at	Thailand'

'He studied in Thailand.'

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (9)
SDM01-20010102-001_9.wav

22)

မိုၵ်	တၵ်	လဝ်	ထီ	ၸေ	ၵုၵ်	တဲ
mün^{2}	tak^{1}	lâu^{5}	thī^{1}	cī^{4}	kon^{2}	tai^{2}
3SG	WILL	tell	close	point	people	Tai

'He would tell many things and point out the Tai people.'

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, No. (54)
SDM01-20010102-001_54.wav

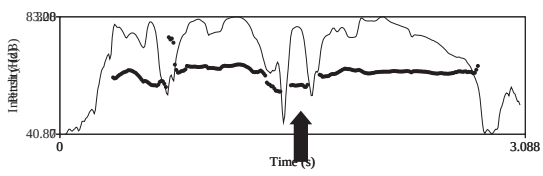
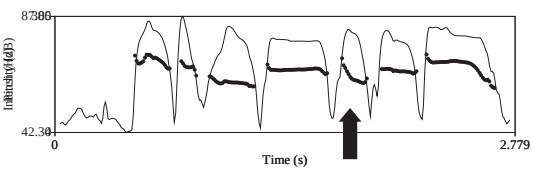
Example (21)	Example (22)
	

Table 16: Demonstrating the realisation of tones in the *Khe Khyang* style (6)

There is no appreciable length difference between the two highlighted elements; if anything the word /tī^5/ in (21) is slightly shorter at around 0.15" than /chī^4/ in (22) which is around 0.18" long. What is clear, however, is that /chī^4/ in (22) shows a pronounced fall, whereas /tī^5/ in (21) does not. One possible explanation for this is that perhaps Banchob's characterisation of the tones (Table 1) is more accurate, as it applies to this data, than ours (Table 2); in other words in this song sung by Aije Let, who is now in his late 70s, the distinction between Tone 4 and Tone 5 is related to contour, rather than duration.

The length contrast can be seen in (19) above, where in the final syllable, /ke^5/ is clearly much longer than the penultimate /phā^4/. The length difference between these two syllables, as we see in Table 15, is much greater than that between the last two syllables of for example (13), shown in Table 9 above. In (19), the 6th syllable is shorter than the 5th (a word with Tone 2, one of the longer tones), whereas in (12) this is reversed.

The shortness of Tone 4 in the rhyming position is demonstrated in (23) and (24). The vowel length of the final syllable, *wai^4*, of (24) is 0.34", in contrast to the much longer final syllable of (24), *yū^1*, which is twice as long around 0.72". The rhymed examples of Tone 4 are both shown with arrows in Table 17 below.

23) မိုၵ် ၸံ ၵုံ ထွၼ် ထုၵ် တီ် ဝဲ
mün^2 cam^2 com^2 phaü^1 thuk^4 tɛm^3 wai^4
he self glad desire copy write keep
'He is glad and interested to write (our language).'

24) ပုၼ် တဲ ၼိုင်ၼဲ ၶံ ယူ
puŋ^1 tai^2 nüŋ^1 nai^4 nam^6 nam^6 yū^1
story Tai such many STAY
'The stories of the Tai a so many.'

Song in the Khe Khyang Style – to honour Stephen Morey, sung by Am Chaw Khya, Nos. (18) and (19)
PhakeSongs_AmSaeuKheKhyang_18and19.wav

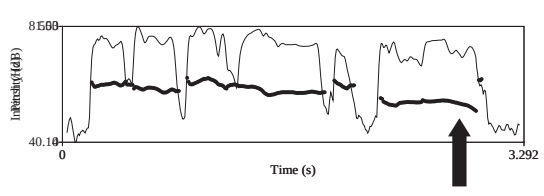
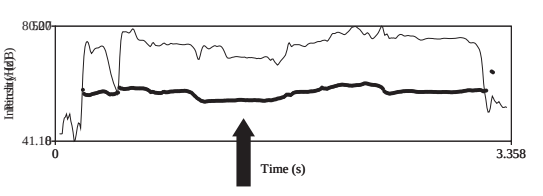
Example (23)	Example (24)
	

Table 17: Demonstrating the realisation of tones in the *Khe Khyang* style (7)

The only conclusion that can be drawn from the data we have presented in this section is that the speech tones of Tai Phake are realised in the *Khe*

Khyang style, in most cases fully represented, although in the case of Tone 3, the phonetic residue of final glottal seems to be the feature that gives the cue to the tonal category.

Is there a melody in *Khe Khyang* independent of the realisation of speech tones?

In this section we will briefly discuss the question of whether we can identify melody that is separate to the realisation of the tones. As we have already seen, in the rhyming positions in Aije Let's *Khe Khyang* (4th and 7th syllables), the pitch of the musical realisation of each token of Tone 1 is always the same – level on the pitch of the 'tonal centre' (cipher notation 1). Similarly the pitch of each token of 2 is always the same, level on a pitch one tone above the 'tonal centre' (cipher notation 2) and then falling.

This leads us to ask the question of whether the *Khe Khyang* style is some kind of musical rendition of speech, without any separate 'musical melody' but following the same pitch patterns as speech would?

We will adduce three pieces of evidence to suggest that there is some kind of underlying melody:

i) The difference between the pitches recorded when several lines were sung first with words and then repeated on a vocable *ta*,
ii) The difference between the pitches recorded when certain words were sung from the pitches when they were spoken, and
iii) Certain patterns where the same tone appears on different pitches in different parts of the song.

In order to discuss this, we will present a narrow pitch transcription of the lines of *Khe Khyang*. For each line of the *Khe Khyang* text, four lines of analysis are presented:

1) Pitch measured in Hz, rounded to the nearest whole number,
   Pitch measured in semitones (where 0 = 440Hz),
   rounded to 2 decimal points (cents),[21]
2) The speech tones of the lexemes (words) in each of the 7 syllables, shown in red, and
3) A transcription of the resultant melody (based on lines 1 and 2), using cipher notation.

To compare the *Khe Khyang* sung with words and with vocables, we present Table 18 which details the pitches found in example (25):

21 The first two lines therefore contain the same information. Some linguists are more used to reading Hz and musicologists semitones and hence we are using both notations.

25) တီ ဟ်ဝ် ထံ ဢော် ၶ ဢီ ပုၼ်
tī⁵ khau⁶ thām⁶ au² khɔ³ lāī⁶ puŋ¹
at 3PL ask take subject many story/meaning
"'Ask the (Tais) and take much information and many stories'".

ၶ ၵၢၵ် ၵ်ဝ် ယုံ ၼဲ ၸွၵ် ၶုၼ်
kɔ⁴ hak⁴ kau² yum¹ nai³ sɔŋ⁶ kon²
friend love 1SG faith get two person
"'I have two dear friends there'".

Song in the Khe Khyang Style (SDM01-20010102-001) sung by Aije Let Hailowng, Nos. (33 and 34)

	σ1	σ2	σ3	σ4	σ5	σ6	σ7
33 sung (Hz)	146 116	114 176	135 183	182 182 114	121 129	134 177	155 157
Semitones	-19.06 -22.95	-22.72 -15.78	-20.35 -15.13	-15.21 -15.24 -23.29	-22.29 -21.14	-20.61 -15.71	-18.05 -17.83
Tones:	**5**	**6**	**6**	**2**	**3**	**6**	**1**
Underlying Melody:	7̣ F	5̣ R	6̣ R	2 F	5̣ R	6̣ R	1
33 vocables (Hz)	151 149	153 154	173 170	176 174 119	129	123 157	153 157
Semitones	-18.44 -18.66	-18.19 -18.14	-16.08 -16.43	-15.76 -15.97 -22.51	-21.16	-22.03 -17.78	-18.24 -17.82
Tones:				**2**			**1**
Underlying Melody:	1	1	2	2 F	6̣	5̣# R	1
34 sung (Hz)	139 119	134 133	181 181 155	156 155	138 109	133 160	187 186 118
Semitones	-19.97 -21.61	-20.50 -20.59	-15.32 -15.33 -18.00	-17.88 -17.95	-19.97 -24.00	-20.66 -17.47	-14.78 -14.83 -22.77
Tones:	**4**	**4**	**2**	**1**	**3**	**6**	**2**
Underlying Melody:	6̣ F	6̣	2 F	1	6̣ F	6̣ R	2 F
34 vocables (Hz)	185 183	177 132	139 133	156	142 136	158	179 179 111
Semitones	-14.97 -15.16	-15.70 -20.82	-19.91 -20.63	-17.88	-19.48 -20.32	-17.70	-15.51 -15.60 -23.74
Tones:				**1**			**2**
Underlying Melody:	2	2 F	6̣ F	1	7̣ F	1	2 F

Table 1 Melodic analysis of Nos. 33 and 34 with Audio Edited.wav
Khe Khyang Song_Nos. 33 and 34 with Audio Edited.wav

As can be seen, the pitch and resultant surface melody in the version with words is quite different from the version sung with vocables. Consider Line (34). When sung with words, the first four syllables produce a melody ḍ ḍ 2 1, whereas when sung with the vocables, the melody was 2 2 ḍ 1. The fact that Aije Let did sing a melody at all when singing vocables suggests that there is some kind of underlying melodic structure, even if its diapason is quite narrow (ḍ 1 2). Our second piece of evidence relates to an example where we can compare the sung text with an almost identical string of words in speech. When making a translation of Aije Let's *Khe Khyang*, we replayed each line to him and he sang it again and then explained the meaning in everyday speech. Example (26) presents such a pair. In both cases the first four words have the same tonal pattern, although the third word is different. The pitch analyses of these first four words in singing and speech are given in Table 19.

26) ꯍꯣꯝ ꯀꯥꯏꯦ ꯃꯨꯟ ꯇꯥꯟ ꯇꯣꯛꯇꯣꯇꯔꯦꯠ
ho⁶ caü² mün² tan² tɔk¹ tɔ¹ rɛt¹
head mind 3SG aim Doctorate
'The aim of his thinking is a doctorate.'

Spoken
ꯍꯣꯝ ꯀꯥꯏꯦ ꯑꯥꯟ ꯇꯥꯟ ꯅꯥꯏ ꯈꯥꯏꯦ ꯅꯥꯏ ꯇꯣꯛꯇꯣꯇꯔꯦꯠ
ho⁶ caü² an² tan² nai⁴ khaü³ nai³ tɔk¹ tɔ¹ rɛt¹
head mind CLF aim this want get Doctorate
'(Your) mind's aim is to get a doctorate.'

Song in the Khe Khyang Style (SDM01-20010102-001)
sung by Aije Let Hailowng, No. (11)

11 sung (Hz)	152 225	193 186 160	186 183 169	185 184 136
Semitones	-18.38 -11.54	-14.23 -14.85 -17.48	-14.87 -15.10 -16.54	-14.99 -15.08 -20.23
Tones:	**6**	**2**	**2**	**2**
Underlying Melody:	ḍ# R	2 F	2 F	2 F
11 spoken (Hz)	146 164	210-182	156 156 147	200-196-154
Semitones	-19.09 -16.99	-12.74 -15.27	-17.85 -17.68 -18.94	-13.65 -14.00 -18.07
Tones:	**6**	**2**	**2**	**2**
Underlying Melody:	ḍ# R	3 F	7F	2# F

Table 19: Melodic analysis of the *Khe Khyang* singing – versus spoken (Line 11)
SDM01-20010102-002_11_1.wav

The noticeable difference here is between the pitches of the three words carrying Tone 2 in the spoken version versus the sung version. In the sung version they carry very similar pitch, but in speech the third syllable *an²* is lower in pitch and quite unstressed. Also, in singing each of the examples of Tone 2 is level before falling, whereas this is not always the case in speech. A second important difference is that in the sung version, the first word (on Tone 6) rises to a pitch well above the commencement pitch of the second word (on Tone 2), whereas this is not the case in the spoken version.

These differences suggest that the rules of speech and the rules of singing are not the same. It may be that the phrasal intonation (which de-stresses *an²*) is not observed in singing, but whatever the cause, this example can be used as evidence that the *Khe Khyang* style is not a simple rendition of speech.

The final example suggestive of some kind of melody independent of tones is that the same Tone in different positions does not always have the same pitch. We have already mentioned that Tone 1 in the rhyming positions (4th and 7th syllables) is realised on a pitch of the tonal centre. However, when there is a Tone 1 on the 4th syllable and a another Tone 1 on the 5th syllable, we notice that the pitch always takes a step up, and the Tone 1 on the 5th syllable is realised a full tone above the tonal centre (i.e. cipher notation 2). This can be seen in Example (15) above, the pitch analysis of which is in the appendix.

While these three pieces of evidence are not perhaps strong enough to posit an 'underlying melody' for the *Khe Khyang* style, they are at least suggestive that there is more happening here than just the 'musical' realisation of speech.

Conclusions

We can conclude from this detailed study of *Khe Khyang* that:

a) tones are realised in the *Khe Khyang* style, especially for words in the rhyming positions (/ta¹ phū³/ and /ta¹ mɛ⁵/);
b) in rhyming position, the creaky tone (Tone 3) is realised with final glottal, whereas in non-rhyming position this may not be present;
c) in rhyming position, the difference between Tone 1 and Tone 5 may be solely one of pitch;
d) in both rhyming and non-rhyming positions, Tones 3 and 4 are clearly shorter in duration in sung style, and this shortness is probably the main clue to identifying the meaning of the particular word;
e) for words in non-rhyming positions (1st, 2nd, 3rd, 5th and 6th syllables), in many cases tone is fully realised;
f) it appears that so long as at least some of the words have their tones fully realised, the meaning of those words that do not realise tones can be found via context;

g) it may be that in singing the tones take forms closer to those found in citation by Banchob in the 1970s, rather than those found in citation by us 30 years later – which seems to particularly apply to Tone 3 and Tone 5, but not to Tone 2; and
h) there is some evidence of a melody that is independent of speech.

Our findings relating to *Khe Khyang* agree with Villepastour (this volume), who wrote of song in Yorùbá that "chanted and sung text deviates from spoken and written Yorùbá in ways that are both formulaic and mimetic of natural speech's own 'deviations' from expected speech tones, rather than serving the aesthetic needs of music, as many scholars have repeatedly claimed." On the other hand, we cannot agree with Schellenberg's (2012) conclusion that "Music is willing to accommodate language as long as this does not interfere with the music."

We suggest that a fuller study of all the different Tai Phake musical styles would be needed to confirm that these findings apply to styles other than *Khe Khyang*. It will be important to examine those styles where much longer line lengths apply and there are consequently fewer rhyming positions. For example the *Kham Phoi* (composed song) can have anywhere between 7 and 41 syllables in a line, according to Aije Let. *Kham Phoi* are also arranged in pairs of lines, of varying length. The /ta^{1} phū^{3}/ in this style is the last content word in the first of the pair of lines, which may be last, 2nd last or even 3rd last syllable of the line; in the second line of the pair there are usually two /ta^{1} mɛ^{5}/, but even with this, the proportion of rhyming positions is much lower than in *Khe Khyang*. One feature of the *Kham Phoi* is the increased use of elaborate expressions, which are themselves a hint to disambiguating meaning.

We also suggest a deeper study of those styles whose musical melody has a wider range and contains some melismas. The *Mo Kham Soi Yoi* (rice pounding song) would appear to be a good candidate for this. In this paper we have offered methodologies for studying the detail of the relationship between speech tone and the musical realisation of spoken texts. This is only a beginning of what must be a much deeper study.

Please visit http://www.phonogrammarchiv.at/7DE709JKSAB
for additional appendices.

References

Bandhumedha, Banchob. 1987. *Phake-Thai-English Dictionary*. Self-publ. [English and Thai]

Chao, Y.R. 1924. "Singing in Chinese". *Le Maitre Phonetique* 39 (April–June): 9–10. [English, written in IPA script]

Gedney, William J. 1972. "A Checklist for Determining Tones in Tai Dialects". In: Smith, M. Estelle (ed.). *Studies in Linguistics in Honor of George L. Trager.* The Hague: Mouton, 423–437.

Li, Fang-kuei. 1977. *A handbook of comparative Tai*. Honolulu: University of Hawaii.

List, George. 1961. "Speech Melody and Song Melody in Central Thailand". *Ethnomusicology* 5: 16–32.

Mendenhall, Stanley T. 1975. "Interaction of Linguistic and Musical Tone in Thai Song". *Selected Reports in Ethnomusicology* II (2): 17–23.

Morey, Stephen. 2005a. *The Tai languages of Assam – a grammar and texts.* Canberra: Pacific Linguistics.

Morey, Stephen. 2005b. "Tonal change in the Tai languages of Northeast India". *Linguistics of the Tibeto-Burman Area* 28 (2): 139–202.

Morey, Stephen. 2010. "The realisation of tones in traditional Tai Phake songs". In: Morey, S. & M. Post (eds.). *North East Indian Linguistics.* Vol. 2. Delhi: Cambridge University Press India, 54–69.

Morey, S.D. & Jürgen Schöpf. 2011. "Tone in speech and singing: a field experiment to research their relation in endangered languages of North East India". *Language Documentation and Description* 10: 37–60.

Schellenberg, Murray. 2012. "Does Language Determine Music in Tone Languages?" *Ethnomusicology* 56 (2): 266–278.

Ware, Vicki-Anne. 2006. *Stylistic and cultural transformations in Bangkok fusion music from 1850 to the present day, leading to the development of Dontri Thai Prayuk*. PhD thesis, Monash University.

The mouth organ *Qeej* as a speech surrogate of the Hmong people in Northern Thailand

(Gretel Schwörer-Kohl)

Introduction

The Hmong are the descendants of an archaic group of people that presumably already in the 3rd millennium B.C. were living in the basin of the Yellow River (Lin 1941: 5). The considerable discourse about the history of the Hmong in the literature during the last few years suggests that in the course of time they were driven away from fertile plains to the mountainous areas of Southern China and to the Northern parts of mainland Southeast Asia.

The Hmong are a subgroup of the Miao people. An estimated number of 12 million Miao live scattered through China, Vietnam, Laos, Thailand, and since the end of the Vietnam War in the diaspora outside Asia.[1]

Of the many divisions among the Miao people in Thailand the two major Hmong groups are: White Hmong (*Hmoob Dawb*) and Green/Blue Hmong (*Hmoob Leeg/Hmoob Ntsuab*), named after the traditional colours in women's dresses.

In China Miao is written with Chinese characters or with an alphabet known as Pollard Miao, in Thailand with the Thai alphabet. In Vietnam the Hmong use the Pahawh Hmong alphabet.[2] At least fourteen major attempts to develop writing systems for the Hmong language can be found over the past one hundred years. Six of these are still in current use (Smalley, Vang & Yang 1990: 149). Today most Hmong write their language with the Romanized Popular Alphabet (RPA), a version of the Latin alphabet developed by Father Yves Bertrais, a lifelong Oblates of Mary Immaculate (OMI) missionary who worked with the Hmong for more than 55 years, and the American missionaries Barney and Smalley. In the RPA tones are indicated by final consonants.

---

1 China 3 million, Vietnam 1,068,189 (2009), Laos 460,000 (2005), United States 260,073 (2010), Thailand 151,080 (2002), France 15,000, Australia 2,190, French Guiana 1,500, Canada 600, Argentina 600, Germany 500, according to more or less recent figures from Wikipedia (s.v. "Hmong people"); different figures can be found in the Wikipedia entry for "Miao people", with 9 to 12 million in China only. Particularly the figures for China vary considerably.

2 See the entries for the keywords "Chinese", "Pollard Miao", "Thai" and "Pahawh Hmong" at <http://www.omniglot.com/>.

For my analysis I will also use the RPA system of writing. This paper is an addition to the existing literature about the transfer of speech tones to musical pitches in the funeral repertoire of the Green Hmong multiple free reed pipe. It summarizes earlier findings of the author and other scholars with both Green and White Hmong in earlier publications (Mareschal 1976; Schwörer-Kohl 1984, 1986, 1991ab, 1997; Catlin 1982, 1986 and Falk 1996, 2003, 2004abc) and confirms the system, while adding another transcription and focusing on poetical language. It also recounts another Hmong story about the origins of the *Qeej* compared to Falk (2004abc).

There is an abundance of speech surrogates among the minorities of Southern China, Northern Thailand, Laos, Myanmar, Vietnam, and Northeast India. Very common are coded messages on jew's harps for entertainment and courting. Flutes are very popular for expressing your sorrows, talking to yourself to solve problems and improvise texts. The author has gathered plenty of material on speech surrogacy in flutes and jew's harps used in courtship, among neighbouring ethnicities, but did not find the time yet for publication.

Comparing all this material, however, we find the most unique speech surrogates among the Hmong, who by the help of their mouth organ *Qeej* have managed to memorize and hand down their very long, elaborate ritual texts for centuries with this remarkable combination of coded texts, instrumental music, and dance steps.

The metre in poetry as well as in music helps not to lose any syllables in the texts. The rhymes help not to change and mix up vocals and tones. The dance steps help to keep the greater structures. Compared to the huge Hmong repertoire very few transcriptions of the music exist so far.

Fieldwork

Concerning field work, ethnographic background and provenance of the research, the author lived in Northern Thailand for two years from 1974–76 and stayed in many Hmong villages. Several funeral ceremonies have been documented in the course of time.

In the beginning, the music for *Qeej* sounded rather monotonous to my Western ears. But I was impressed by the length of the compositions and the seriousness and concentration of the players. How complex the forms of the pieces are I only learned much later, when analysing the different parts of these compositions. Some pieces showed surprisingly "classical" forms (Schwörer-Kohl 1997). During the funeral ceremonies the coded texts will never be heard.

I became aware of the connection of music and language only after I had asked a *Qeej* player to teach me at least one simple piece on this fascinating

instrument. I was initiated as a student, after I had brought the usual offerings (a bottle of white liquor and some other gifts). Despite being female, I seemed to have been considered a strange neutral person and was allowed to touch the instrument, although by that time in my area of studies Hmong women were not at all allowed to play the *Qeej*. Very soon, when still practising how to breathe properly, the teacher explained to me that I could only become a good player if I learned the Hmong language – which made me curious. So I discovered this phenomenon not in a refugee camp, where I recorded a lot of music in the following years, but in the Hmong village of Mae Sa Mai, today only half an hour away from Chiang Mai. In the seventies it took about two hours on bumpy roads to get there. In the beginning, it was very hard to find someone to explain the texts, as they were considered to be secret and full of magic power. Recordings and explanations were only possible outside the village, when teachers and students were practising together, usually in the evenings after dinner. The players were always careful not to attract the wrong spirits. In the summer of 1975, after thousands of refugees had come from Laos to Thailand, it was much easier to get information in the camps. But even though the spirit world was considered to be disturbed, the players would still look for special places, not to attract bad spirits to the wrong places. The men had plenty of time. I did not have to wait until they came back from the field work, and there was a strong will to hand down the music and other traditions. That was also the time when people of different Hmong clans started to write down the texts themselves. Finally thick mimeographed books would circulate. Now comparative studies with the music of the Hmong in China are a desideratum. The question will be how many of the traditions have survived the cultural revolution. So far during my recent travels, I have only noticed that the intimate festivities and rituals in Northern Thailand can hardly be compared to the huge organized festivals in Southern China. Judging from the patterns of the embroideries, the Green Hmong of Northern Thailand seem to come from Guizhou. Further studies (as to whether the musical styles in these areas are related to each other, too) are also desirable.

The mouth organ *Qeej*

Now a few words on the *Qeej*, the most important musical instrument of the Hmong people and the most interesting one for speech surrogates. The etymology of its name is unknown. It consists of a wind-chest of hardwood into which six bamboo pipes are inserted. The upper end of the air-chamber serves as a mouthpiece. One of the six pipes, the *ntiv luav qeej*, is considered as the "leading pipe of the mouth organ" (Bertrais 1964: 150) and has three times the diameter of the other bamboo canes and three free reed tongues. Into each of the five thin pipes only one single tongue is inlaid.

The following Hmong story, which I recorded on January 9, 1981 in the refugee camp of Ban Thong close to Chiang Khong in Northeastern Thailand, provides valuable information about the mouth organ as a language substitute. Txij Choj Yaaj, who had fled from Laos five years earlier, told the following legend (summarized):

> A long time ago, our ancestors lived in China. As they did not have their own land to dwell on, the Chinese after some time banished them. So they took refuge in the direction of the Chinese border. By that time they kept a book in which, on a long piece of cloth, instructions were written down what to do when a person had died, how to send the soul to the huge *ceeb tsheej* (realm for the souls of the ancestors). On their migration the Hmong people pulled the cows carrying their belongings with ropes. When they had reached the place called Yeej Yaj-market, they decided to have a rest for one night. Here they noticed that their book was very wet from rain. So they untied and displayed it in the sun to get dry. Nobody was watching the cows that were very hungry. So two of the animals approached the book and started to eat it. When our ancestors noticed the disaster, only a small part was left. This fragment they put into a basket. During a stop in another village, where they wanted to have a look at the remains of the book, they noticed that the rats had eaten the rest. In desperate need of help they decided to see *puj saub siv yig* (female helpful power in Hmong religion). She gave the following advice: "One person has to go to cut six bamboo canes for six pipes, one person has to melt metal to form the free reeds and one person has to hurry up, to cut wood for the wind-chest of the mouth organ. Whenever a person has died, you use this *Qeej* to explain him the way to his ancestors. The *Qeej* player will be the one who can send the deceased, as the book has gone. He will be the only one, who will know how to lead." After the Hmong had made the mouth organ, they asked *puj saub siv yig* how to play it. She said: "You are six brothers. Each person plays one pipe, so six brothers play six pipes to guide the spirit of the deceased. In this way the six brothers blew their instruments for many years. When all of them had died, the descendants again did not know how to see off the dead people. So they had to consult *puj saub siv yig* again. She answered: "As six men are no longer able to play, take a piece of wood and bore holes through it. Then insert six pipes. So six fingers of one person will be able to blow the 6 pipes of the *Qeej*." Thus *puj saub siv yig* has taught the Hmong people that when a human being has died, the *Qeej* has to be played to explain him or her the way to *ceeb tsheej*, to the grandfathers' and grandmothers' old country in the great land of China.

Concerning the symbolic connotations of this legend we learn that the *Qeej* is a book substitute and – for the soul of the deceased – the leader to the other world, to the country of the ancestors. These two functions give this musical instrument a very high reputation in Hmong society. As these tribal people for several centuries have been living on the margin of refined, civilized people, they adore the ability to read and write. Because of their own illiteracy they

suffer from inferiority feelings. However, it fills them with pride and self-confidence that they are able to compensate this deficiency with their very own, unique way of handing down their tribal knowledge: they code texts on the mouth organ. The melody follows the tones of the language and some motifs express special word combinations. This coded language has many advantages: it offers the possibility to communicate with the souls or spirits of the deceased people as well as the sacrificed animals. Moreover, melody and rhythm help to memorize and hand down the ritual texts and the other forms of poetry. Even in the nomenclature for the six pipes of the *Qeej* we find connotations related to the task of the mouth organ as a speech surrogate: while four pipes have onomatopoeic names or designations that point to their local position, the thick pipe is called *ntiv luav* (*ntiv* = 'pipe'; *luav* = 'leader' [in this context]). Not only because of its dominance in outer appearance but also in music, where it sounds more loudly and frequently than all other pipes, is it considered as the leading speaker. Furthermore, it is the guide for the soul of the deceased to the world of the ancestors. The pipe behind the "leader" is called *ntiv laig* (*laig* = 'bringing offerings to the souls of the ancestors'). Its name points to the ritual function of the *Qeej* playing and reveals an additional aspect: the coded language is not only very useful for the communication with the spirit world and for handing down traditional texts, but also as an offering to the ancestors, as they are believed to enjoy this music with great delight. Among the "speaking pipes", this "offering pipe" is the most eloquent one.

The Hmong funeral repertoire

The Hmong funeral ritual – traditionally a three-day and night ceremony – is the most important affirmation and reinforcement of Hmong identity for Hmong people in the original homelands in China and in the diaspora, the mountainous regions of Southeast Asia and the Western countries. The *Qeej's* coded messages to the soul of the deceased are not only of interest to ethnomusicologists and linguistics as one of the more complex examples of speech coding in music but also anthropologically – *why* are these important texts so concealed? The repertoire of compositions for distinct stages of the funeral rituals is quite huge. According to 50-year-old Vaag Chao Sae Xiong, who fled from Laos in 1980 and also played the part that is transcribed here, the most important pieces are:

1) *Qeej tu sav* (*Qeej*, when the breath has been cut off)
2) *Qeej tsaa neeg* (*Qeej*, when riding the horse) played while the corpse is laid on a stretcher, serving as his horse and carrying him part of his way to the ancestors
3) *Qeej noj tshais* (*Qeej*, while eating breakfast)

4) *Qeej noj su* (*Qeej*, while eating lunch)
5) *Qeej noj hmo* (*Qeej*, while eating dinner)
6) *Qeej hlawv ntawv* (*Qeej*, while burning paper money) when the rice-paper money is transformed into ashes, it can be used by the dead person's spirit to defray the expenses on the way to the ancestors
7) *Qeej sawv kev* (*Qeej*, when sending off) played, when the corpse is brought to the burial place.

Parts 3), 4) and 5) accompany the respective meals that are offered to the deceased. Besides these compulsory compositions, the Hmong play additional pieces for fun and for the diversion of the deceased as well as the bereaved. Thus during the ceremonies the *Qeej* will be heard day and night with only a few hours of intermission. The sound may penetrate through the whole village and even reach the surrounding valleys. The *Qeej* can also be blown if no case of death has been reported. The Hmong hold a huge repertoire of compositions "for fun". Their coded texts are based on tribal legends.

The composition *Qeej tu sav*

How the coding of syllables works will be explained in a part of the composition *Qeej tu sav* (= mouth organ about cutting off the breath) called *ntxuav muag* (= 'washing the face'; literally: 'the eyes'):[3] The composition *Qeej tu sav* with all its repetitions can last for more than an hour. In the part analyzed here, the *Qeej* player explains to the dead person that he has to be washed to get prepared for his long journey.

In transcription 6 the part of the mouth organ player is noted down. Under the chords the syllables which the player has in mind during his performance are written down. In transcription 6a we have the sung melody that the player had to learn from his teacher by heart before he was able to play the corresponding *Qeej* version. The transliteration of the Hmong syllables is complicated and looks puzzling, as sometimes too many consonants follow each other. But it has the advantage that phonetically it is very accurate. Except for nasals, shown by doubling the vocals that have to be nasalized, the Hmong language does not have final consonants. So consonants can be used as tone marks, like B J V S G M. Green Hmong has 7 tones. Syllables in the texts that end with a single vocal are spoken on the neutral, middle tone. The singing follows the speech tone. The *Qeej* melody adapts the song melody and so indirectly also follows the speech tones.

3 For a transcription of the part *tu zaam* (= 'dressing the corpse') see Cooper, Tapp, Lee & Schwoerer-Kohl (1991: 45); for explanations to the last part *zais tsuj zais neev* (= 'concealing trace of return') see Schwörer-Kohl (1982: 609–617).

Here is a complete version of the Hmong text including all sung as well as blown syllables (the syllables that appear only in the sung version and are not coded in the *Qeej* part have been underlined; the syllables that appear only in the instrumental and not in the sung version are in brackets):

Taav					
Nuav ab nraug lis vaug yuav tuag lis tag moog tuag deev					su loo
	yuav moog tuag tu sav		qos zim tuag ntawm ncauj tuag tu paa		
	qos zim tuag ntawm lub				su
(yuav) ab nraug lis vaug yuav leeg kwv kws leej teg yuav dej txag			qos zag yuav txawm hov twg li tuaj dej txag		
			qos zag yuav txawm qaab cij		(loo)
(yuav) zeb		oob txag	qos zim yuav txawm hov twg li lug		
(yuav zeb)		oob txag	qos zim yuav txawm qaab cij	ca	su (loo)
yuav ab nraug lis vaug (yuav) leej kwv kws					
leej teg yuav moog muab		dej txag	qos zim yuav lug rhaub ua dej		su
				(ca)	su(loo)
ntxuav ab nraug lis vaug	cev quav luaj dub (yuav) nrum		qos nroog tej qaab cub (yuav)		
ntxuav ab nraug lis vaug	cev quav luaj daag(yuav) nrum		qos nroog tej qaab txaag		
	(yuav)	dej txag	qos zag moog ntxuav taag		
	(yuav)	oob txag	qos zim moog ntxuav txhua		su loo)
			ntxuav tau		(su loo)
ab nraug lis vaug	(yuav) tug ntsuj maag		qos plig dawb (yuav) moog		
			qos nyoog yuav moog cuag puj		
			qos tag moog cuag yawm	ca	su loo

This is the translation of the coded syllables:
Now the gentle person (addressing the deceased) is dead, is indeed dead. *Su Loo*. He has died. He has stopped breathing. In his mouth the breath was cut off. He is dead and can utter not one single word more.
The younger brother, the elder brother of the gentle person will get cold water in which place?
They will get cold water in the forest.
The brothers and cousins of the gentle person will go to get cold water.
They will heat it up to become warm and then wash the body of the gentle person, the black heap.
Then they will throw the water to the floor between the wall of the house and the fire place.
They wash the body of the gentle person, the yellow heap.
Then they will throw the water to the floor under the bed.
They go to have the washing with cold water, when they have cleaned all parts, when they have finished the washing, the spirit of the gentle person will be clean and soft and will start his journey, to meet the grandmothers and the grandfathers.

The table below in the first row on top shows the letters for the tone marks and underneath the level on which the tones are spoken.

In the second row we see how the speech tones sound in melodies, and in row 3 we have the corresponding chords, the sounds, or clusters on the mouth organ.

Row 4 shows some typical melodic figures. The first one, the address to the dead person, will only appear in funeral music. The other coded syllables can also be heard in the compositions for fun and pure entertainment.

taav nuav (= 'now') announces the beginning of a new part. *Su loo* (= 'for sure, certainly') announces the end of a part; as a recurrent formula it could be compared to the Christian "Amen". The verb *yuav* (= 'will, have to, should') gives advice to the dead person. This verb appears where it has to stand grammatically, but its second function is to give the player the opportunity to inhale air for the next passages. So *yuav* can be interspersed on many occasions. *Ab Nraug Lis Vaug* is a polite soft name for the deceased. This melodic and rhythmic formula is a kind of leitmotif in the funeral songs of this clan among the Green Hmong. It makes listeners immediately aware that they hear a funeral composition. The texts for the compositions are not in everyday language, but in a sophisticated poetical form, rather difficult to translate even for eloquent experienced old Hmong men. As stylistic elements we have questions and answers. Typical for Hmong rhetoric are also word pairs with a combination from different language levels such as the archaic poetical, religious language and idioms from modern everyday language. The syllables have mnemotechnic, onomatopoetic and magic functions.

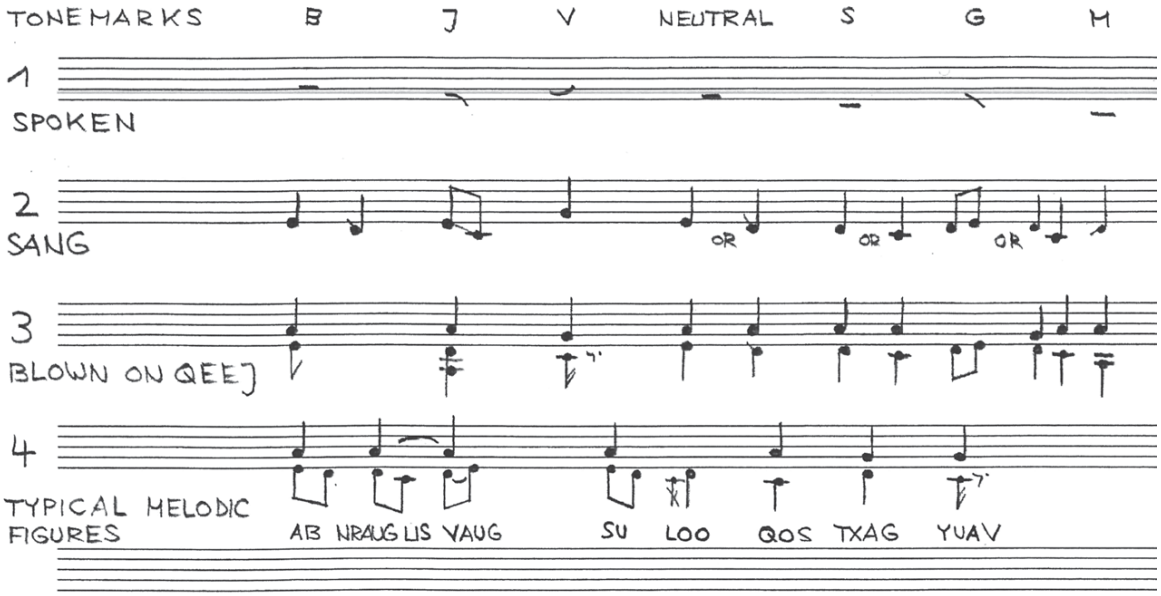
Levels of understanding *Qeej* playing

The question should be raised to which degree not only the spirit of the deceased person will understand the text but also the living listeners. Men who play *Qeej* themselves will understand very well, men who know a lot about religion also understand pretty well. Many women do not understand very much as traditionally they do not learn to play the *Qeej* and are not allowed to perform most of the religious ceremonies.

What will help listeners to understand the coded texts? Here we have part of a composition in which each syllable has a semantic meaning. But there are also other sections with many onomatopoetic syllables that are typical for specific parts of a composition. So a newcomer will first find out which part of the long composition is played. The praeludium *ntiv* for example has characteristic formulas etc. Then he will wait for leitmotifs addressing a dead person, a special spirit or the formula "the rats have eaten all", the latter announcing that a story is told for entertainment and that it is not a composition for spirit worship that is being performed.

For the *Qeej* playing transcribed here the experienced listener will hear the formula *tav nuav*, which will tell him that a small section of the main part is played. The melody for *Ab Nraug Lis Vaug* will make him aware of a funeral song. After many syllables repeated in all the small sections he will watch for the rhetorically important questions and answers in which the essential information is imbedded. These will give him more orientation and make him understand the texts.

An old, very good mouth organ player once characterized the highly complex relation between tonal language and mouth organ playing in the following way: while I blow the mouth organ I sing in my heart.



Please visit http://www.phonogrammarchiv.at/7DE709JKSAC
for additional appendices.

References

Bertrais, Yves. 1964. *Dictionnaire Hmong-Francais*. Vientiane, Laos: Mission Catholique.

Catlin, Amy. 1982. "Speech Surrogate Systems of the Hmong: From Singing Voices to Talking Reeds". In: Downing, Bruce T. & Douglas P. Olney (eds.). *The Hmong in the West: Observations and Reports*. Minneapolis: University of Minnesota, Southeast Asian Refugee Studies Project, 170–200.

Catlin, Amy. 1986. "The Hmong and their Music". In: Cubbs, Joanne (ed.). *Hmong Art: Tradition and Change*. Sheboygan, Wisconsin: The John Michael Kohler Arts Centre, 10–19.

Cooper, Robert, Nicholas Tapp, Gary Yia Lee & Gretel Schwoerer-Kohl. 1991. *The Hmong*. Bangkok: Artasia Press.

Falk, Catherine. 1996. "Upon Meeting the Ancestors: the Hmong Funeral Ritual in Asia andAustralia". *Hmong Studies Journal* 1(1): 1–15.

Falk, Catherine. 2003. "'If you have good knowledge, close it well tight': concealed and framed meaning in the funeral music of the Hmong *qeej*". *British Journal of Ethnomusicology* 12 (2): 1–33.

Falk, Catherine. 2004a. "The Dragon Taught Us: Hmong Stories about the Origin of the Free Reed Pipes *Qeej*". *Journal of Asian Music* 35 (1): 17–56.

Falk, Catherine. 2004b. "Hmong Instructions to the Dead: What the Mouth Organ *Qeej* Says (Part One)". *Asian Folklore Studies* 63 (1): 1–29.

Falk, Catherine. 2004c. "Hmong Instructions to the Dead: What the Mouth Organ *Qeej* Says (Part Two)". *Asian Folklore Studies* 63 (2): 167–220.

Lemoine, Jacques. 1972. "L'Initiation du mort chez les Hmong". *L'Homme* 12 (1): 105–134, 12 (2): 85–125, 12 (3): 84–110.

Lin, Yueh-Hwa. 1941. "The Miao-Man Peoples of Kweichow". *Harvard Journal of Asiatic Studies* 5 (3/4): 261–345.

Mareschal, Eric. 1976. *La Musique des Hmong*. Paris: Musée Guimet.

Mottin, Jean. 1982. *Allons faire le tour du ciel et de la terre*. Bangkok: White Lotus.

Ratliff, Martha. 2010. *Hmong-Mien Language History*. Canberra: Pacific Linguistics.

Schwörer-Kohl, Gretel. 1982. *Die Mundorgel bei den Lahu in Nord-Thailand: Bauweise, Funktion und Musik.* Hamburg: Wagner.

Schwörer-Kohl, Gretel. 1984. "Sprachgebundene Mundorgelmusik zum Totenritual bei den Hmong in Nordthailand und Laos". In: Mahling, Christoph-Hellmut & Sigrid Wiesmann (eds.). *Bericht über den Internationalen Musikwissenschaftlichen Kongress der Gesellschaft für Musikforschung in Bayreuth 1981*. Kassel: Bärenreiter, 609–617.

Schwörer-Kohl, Gretel. 1986. *Anthologie südostasiatischer Musik: Musik der LA HU NYI in Nordthailand. Aufnahmen, Photographien und Kommentar in Deutsch und Englisch*. Kassel: Bärenreiter.

Schwörer-Kohl, Gretel. 1991a. "Über den Begriff NTIV in der Musikterminologie der Hmong in Nordthailand und Laos". *Jahrbuch für musikalische Volks- und Völkerkunde* 14: 68–96.

Schwörer-Kohl, Gretel. 1991b. "Mouth organ and drum: the symbols of death among the Hmong in northern Thailand". In: Tokumaru, Yosihiko (ed.). *Tradition and its Future in Music*. Tokyo: Mita Press, 249–253.

Schwörer-Kohl, Gretel. 1997. "Zum Werkbegriff in der Ethnomusikologie und in der historischen Musikwissenschaft". In: Mahling, Christoph-Hellmut & Stephan Münch (eds.). *Ethnomusikologie und historische Musikwissenschaft: Gemeinsame Ziele, gleiche Methoden? Erich Stockmann zum 70. Geburtstag*. Tutzing: Schneider, 313–324.

Smalley, William A., Chia Koua Vang & Gnia Yee Yang. 1990. *Mother of Writing: The Origin and Development of a Hmong Messianic Script.* Chicago & London: The University of Chicago Press.

Symonds, Patricia V. 1991. *Cosmology and the Cycle of Life: Hmong Views of Birth, Death and Gender in a Mountain Village in Northern Thailand.* PhD thesis, Brown University.

Xiong, Lang, Joua Xiong & Nao Leng Xiong (comp.). 1983. *English–Mong–English Dictionary*. Milwaukee: Xiong.

Preliminary research into the influence of the process of musical transmission on the relationship between speech tone and melody in Kam singing traditions from southwestern China[1]

(Catherine Ingram)

Introduction

The centuries-old songs of Kam people, a southwestern Chinese minority group known in Chinese as *Dongzu* 侗族, have traditionally been transmitted from one generation to the next through a specific two-stage process. This process of musical transmission involves learning to chant the lyrics of a song and then learning to sing the song, and has not been reported by other researchers. During the 24 months of fieldwork I have undertaken in rural Kam areas since 2004, I have often been present as Kam village singers engaged in these two stages of song learning, and I have audio-visually documented the process. As my skills in the Kam language and my knowledge of Kam musical culture have increased, I have also begun to learn to sing Kam songs through the same method. In this paper, I present a preliminary analysis of how this process of musical transmission appears to influence the complex and largely undocumented relationships between speech tone and melody (hereafter, the STM relationship) in Kam singing.

My research in this area is complicated by two factors. Firstly, and most significantly, there is no comprehensive analysis of the Kam dialect used in singing the songs discussed in this paper – such as the UNESCO-recognized "big song" genre.[2] Secondly, as the first researcher to document the process of musical transmission within Kam singing, I have had to devise my own framework for understanding how the transmission process operates in relation to the STM relationship. Although the two factors that complicate this research are also present in similar ways in other linguistic and musical contexts, there do not appear to be well-established research models for

1 I gratefully acknowledge many Kam singers and song experts for assistance with my research; especially, I thank song experts Wu Meifang, Wu Pinxian, Wu Xuegui and Wu Zhicheng and other singers in Sheeam. I acknowledge financial and other support for this research provided by an Endeavour Australia Cheung Kong Research Fellowship, the University of Melbourne, PARADISEC (<http://www.paradisec.org.au>), the Research Institute of Ritual Music in China (Shanghai Conservatorium of Music) and a Newton International Fellowship (British Academy). Many thanks also to Stephen Morey and two anonymous reviewers for helpful comments on an earlier draft of the written version of the paper.

2 See *Grand song of the Dong ethnic group* at <http://www.unesco.org/culture/ich/index.php? RL=00202>.

dealing with such problems.[3] As a consequence, even preliminary research such as the present study may be of great use to other scholars.

Kam people and the Kam language[4]

Kam people are officially recognized as one of China's 55 minority groups, and are mainly resident in a small area of southwestern China comprising parts of southeastern Guizhou Province and the adjoining border regions of Hunan Province and Guangxi Zhuang Autonomous Region. Today, the Kam population is estimated to be at least 3 million people.[5] Many Kam people continue to speak Kam as their first language, and contemporary forms of distinctive Kam cultural and spiritual/religious traditions continue to be practised in some rural Kam areas.

The Kam language is a tonal, monosyllabic language with no widely used written form, and is classified within the Kam-Sui branch of the Tai-Kadai language family. Significant English-language studies of this language family and, particularly, the classification and features of Kam include Diller, Edmondson & Luo (2008), Long & Zheng (1998), Edmondson & Solnit (1988), Sagart (2004), Matthews (2006) and Gerner's (2009: 737) succ inct overview of the controversies regarding the categorization of Kam with respect to Kam-Tai, Tai-Kadai, and Austronesian languages. The Kam language has two major dialects – Northern Kam and Southern Kam – and each dialect contains three or four lects (that is, regional speech varieties). In the context of daily speech the differences amongst both the local variants and the lects are often minor, whilst the differences between dialects are frequently significant. However, even when similarities amongst variants or lects are obvious, song lyrics are rarely understood in full outside the region from which they originate.

My research has primarily been undertaken amongst communities speaking the second lect of Southern Kam. Today, such communities are those most active in maintaining traditional Kam musical activities. Furthermore, the Kam genre known as "big song" that is inscribed on UNESCO's Representative List of the Intangible Cultural Heritage of Humanity is predominantly sung using the second lect of Southern Kam.

---

3 For example, Lundström (2010) and Miller (1985) discuss both the STM relationship and musical transmission in other southeast Asian linguistic contexts, but unlike the present study are able to rely upon a comprehensive linguistic analysis of the tonal systems of each language in order to draw their conclusions.

4 I gratefully acknowledge the assistance of Stephen Morey, Ruth Singer and Myfany Turpin in preparing this description of the Kam language.

5 As deduced from data on the Kam population in Guizhou recorded in the 2005 census and available at <http://www.chinadataonline.org> (accessed 9 September 2011).

For these reasons I have focussed on this region, and mainly the well-known big-song-singing region that is known in Kam as Sheeam (in Chinese, Sanlong 三龙) and is located in Liping County, Southeastern Guizhou Province.

The second lect of Southern Kam is barely documented in the current literature, and Pan Yongrong's brief study (Pan 2005) is the only publication specifically dealing with the Sheeam variant. In his brief article, Pan states that the variant of Kam that is spoken in Sheeam has 44 consonants occurring word initially (that is, initials or onsets, see Figure 1), 8 vowels (Figure 2), 59 finals/rhymes/syllable endings (Figure 3), and 10 tones (Figure 4).

p	ph	m	m̥	f	w	
t	th	n	n̥		l	l̥
ts*	tsh*			s		
ʦ	ʦh	ɳ	ɳ̊	ɕ	j	
k	kh	ŋ	ŋ̊	x	ɤ	
q	qh					
ʔ				h		
pj	phj	mj			wj	
tj	thj	nj	n̥j		lj	l̥j
kw	khw	ŋw				

Figure 1: The 44 consonants that Pan (2005) states can occur word initially in the variant of one village in Sheeam, presented in his arrangement. Besides the consonants listed here, I find that a small number of words also begin with a vowel; based on Liang & Zhang's (1996: 4) conclusions, some may also begin with a glottal plosive.

i	e	a	o	u	ɐ	ə	ɔ

Figure 2: Vowels that Pan (2005) observes in the variant of one village in Sheeam (deduced from his analysis).

5 Single Vowel Endings:

i	e	a	o	u

10 Composite Endings or Diphthongs:

ai	ɐi	əi	oi	ui	au
ɐu	əu	eu	iu		

44 Consonant Endings:

am	an	aŋ	ap	at	ak
ɐm	ɐn	ɐŋ	ɐp	ɐt	ɐk
em	en	eŋ	ep	et	ek
əm	ən	əŋ	əp	ət	ək
im	in	iŋ	ip	it	ik
om	on	oŋ	op	ot	ok
					ɔk
um	un	uŋ	up	ut	uk
			UP		

Figure 3: The 59 syllable rhymes that Pan (2005) observes in the variant of one village in Sheeam, presented in his arrangement.

Type of tone	Live (appearing in words with open syllables)						Dead (appearing in words with a coda, i.e. with closed syllables)			
Tone number	1	2	3	4	5	6	7	8	9	10
Tonal shape	42	232	55	12	53	21	42	232	55	12

Figure 4: Pan's (2005) analysis of tones in the Kam language in the variant of one village in Sheeam, analyzed using Chao Yuanren's 1–5 degree pattern. In this scheme, 1 indicates the lowest pitch and 5 the highest, with the intervals between these extremes being roughly equivalent.

Overall, the data provided by Pan provides a helpful introduction to the Kam language as spoken in Jai Lao, one of the two large villages in Sheeam. However, Pan's article is not sufficiently detailed or free from obvious inaccuracies (particularly regarding vowel length and the tonal system, as

detailed in Ingram 2010) to be useful as a foundation for more detailed analysis of the STM relationship. Consequently, in this article I am unable to indicate tone in my transcriptions of Kam words as they are spoken in Sheeam. Because of the lack of data regarding the Sheeam variant, and because of the differences between this variant and the lect used to develop the official Kam orthography (promulgated in 1958 and based on the first lect of Southern Kam), Kam words in the main text of this article are transcribed using my own practical phonemic orthography, based on the Roman alphabet and commonly accepted (Australian) English pronunciation. The International Phonetic Alphabet (IPA) is used for song lyrics.

Scope and important features of Kam singing traditions

In many respects, village life in Sheeam and many neighbouring Kam regions continues to be inseparable from and significantly shaped through music. For example, singing takes place at weddings and engagements, after the building of a new house, or on the arrival of important visitors. Praising important guests with suitable songs is still just as important as giving them food and drink. During the procession and other ritual activities carried out in villages in Sheeam and many other Kam areas at New Year, singing is still used to ask for blessings from the female deity *Sa* for the oncoming year. Every day for over a week during the festive season, Kam opera performances or *yeh* singing take place. And in villages in Sheeam and nearby Southern Kam areas, in the evenings during New Year celebrations singing groups gather in the tall pagoda-shaped *dare low*, the impressive wooden tower built in many Kam villages, to carry out a sung exchange of the multi-part choral songs usually known in English as big song.[6]

Kam people classify their song genres in different ways: *ga bai-jin* (Going to the mountain songs), *ga bu-ngan* (Songs praising a new house) and *ga bu-nong* (Songs praising a newly born child) are classified according to the location, activity or occasion with which they are associated. Other genres are classified according to particular musical features (such as the accompanying instruments used for *ga beeba* (*Beeba*/Kam pipa[7] songs)), lyrical content (such as *ga lyen* (Guessing songs) and *ga yeh* (*Yeh* songs)) and historical background (such as *ga nya* (River songs, from the Fulu river region)). Genres such as *ga bu-nong* are sung only by women, and genres such as *ga jin* (Narrative songs) are sung only by men. Accompanying instruments such as *beeba* (Kam pipa) and *gor-gee* (two-string fiddle) are only used for certain

6 Other village locations were sometimes also used for big song performances. See Ingram (2007; 2010; 2011; 2012) for a fuller description of big song and the other Kam musical traditions described in this section.

7 The *beeba*, or Kam pipa, is a four-string plucked lute of varying sizes; see Ingram & Wu (forthcoming).

genres. The composition of the performance also varies between genres and includes solo singing (as for *ga bai-jin* and *ga beeba)*; unison choral singing (as for *ga kwaow* (Wine songs)); multi-part choral singing (as for big song); sung dialogue/song exchange (as for *ga lyen)*; and call-and-reiteration (as for *ga yeh*, wherein the lyrics are sung by one or two soloists and then repeated by the whole group using a slightly different melody). Kam people with some formal musical training occasionally identify the tuning of instruments according to the sol-fa scale. However, I never heard Kam singers explicitly identify the tuning systems or pitch sets used in Kam musical traditions.

As noted in Ingram (2010) the issue of pitching within Kam singing is complex, and without further detailed investigation of this issue I am thus reluctant to ascribe particular pitch sets or modes to particular Kam musical genres. However, by way of a general indication, Figure 5 offers the reader a general guide to the main pitch set used for all big songs in the Sheeam repertoire. The asterisk marks the central pitch that functions as a "tonal centre". Incidental appearances of the half-sharpened fourth-degree and of the tonal centre in the upper octave are omitted here. The pitch set is nominally centred on A3; singers only use a predetermined fixed pitch when performing with an accompanying instrument.

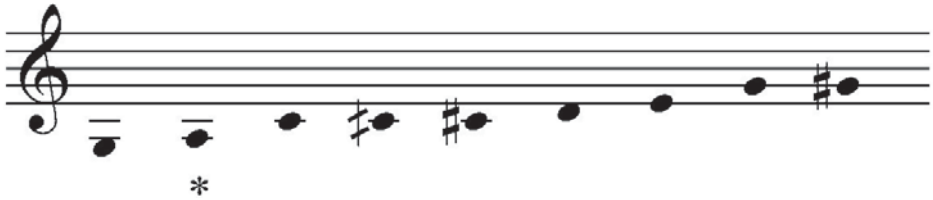


Figure 5: The main pitch set used for all big songs in the Sheeam repertoire.

Many Kam song genres and categories are also distinguished according to their *sor* or "melodic habitus".[8] The Kam word *sor* is usually used to refer to song melodies, and in most cases all songs in the one genre or one category of a genre are said to share the same *sor*. *Sor* is a polysemous word that encompasses a range of meanings from 'voice' and 'sound' to 'breath', 'life' and 'life-force' (in some senses equivalent to the Chinese *qi* 气). Although a certain group of songs may share one *sor*, their sung melodies are not necessarily identical – in fact, there can be melodic differences in the ways that the same *sor* is rendered in each song, in performances of the same song by different groups and, to a lesser degree, in each group's performances of the one song at different times. Since it is clear that Kam singers do not perceive melodies as entities deriving from a single static, unchanging form, it would be somewhat misleading to assume that *sor* are conceptualized as set melodies that are varied for each set

8 In this article I use "category" to refer to a subset of a "genre".

of lyrics. Rather, each *sor* is perceived as a "melodic habitus" (drawing upon the terminology of Bourdieu 1977 [1972]), and the melodic choices for all songs sharing the same *sor* are loosely defined by that habitus.[9] The relatively conspicuous differences in the renditions of one *sor* for different songs often result from differences in the song lyrics, which are also the most evident (and sometimes the only) feature distinguishing different songs within a category.

The prominent role of song lyrics as one of the most evident differences between different songs sharing the same *sor* attests to the lyrics' central importance. This importance is also implicit in the expression *lak ga* – literally, 'the bones of songs' – that is used in the Kam language to refer to song lyrics. Lyrics are very important in judging a song's quality,[10] and can sometimes be interchangeable between genres. Although the same song genre and/or category of songs within a genre, and sometimes the same sets of lyrics, are recognized or used in different Kam areas, the *sor* used for each category or genre of songs in each region or village is unique.

The lyrics to all Kam songs are pre-composed, and almost always involve rhyming patterns. These rhyming patterns are often complex and are considered by most experienced singers and song experts[11] to be integral to the definition of "song". The rhymes used in many Kam song lyrics involve the final sound[12] of syllables at particular points within each line and section. While rhymes involving linguistic tone are well recognized in research on Tai-Kadai poetics,[13] my research indicates that details regarding tonal rhymes are not yet fully explored.

The only rhyming pattern explicitly referred to by Kam people in Sheeam is the rhyme known as *hwa* that usually occurs between the final sounds of the last syllables of either even-numbered lines or whole verses, and which I refer to as an "ending rhyme". The two other types of rhymes common in Kam song lyrics I refer to as "internal rhymes".[14] Figure 6 gives an example of all the

9 Adherence to a "melodic habitus" is a phenomenon apparent in many musical genres both within and beyond China, especially within non-literate cultures. Scholars have not analysed this in a single way, or used consistent terminology (see, for example, Schimmelpenninck 1997; Mittler 2001: 314; Thrasher 1990: xiv, xvii; and further discussion in Ingram 2010, 2012).

10 See also Ingram (2011) for a discussion of the significance of Kam song lyrics in relation to local knowledge of the environment.

11 *Sang ga* (song experts) are the people most responsible for transmitting songs to the next generation, and are considered most knowledgeable about Kam songs.

12 Namely, what linguists call the "rhyme"—the nucleus and coda of a syllable, or the syllable part following the onset.

13 For analyses of Kam poetics, see Yang (1986 [1982]), Yang & Zheng (1986), and Zhang & Liu (1997). Kam poetics share many similarities with those used in other Tai-Kadai languages (see, for example, Gedney 1997; Huang 1987; Hudak 2008; Holm 2003: 32–37).

14 I only once heard these rhymes named in Kam by a song expert, and as I am not yet certain if such names are widely recognized I do not discuss them here. Chinese researchers have described these rhymes in

different types of rhymes that may occur in one set of song lyrics – here, within the lines of the song *Lao nyin mang* (‘Old people happy’, see Appendix 1 for a translation). In this transcription the internal rhymes that occur from an odd-numbered line to the following even-numbered line are shaded, and internal rhymes within a line of lyrics appear below in thin-lined boxes. Ending rhymes are marked with bold script inside thick-lined boxes. As this example shows, the position of the internal rhymes may vary throughout the song, as may the sound on which the rhyme is based.

line													
1	hwan	hwan	ʈon	kha	jau	to	mɐi	ka	sai	ɕau	lau	njin	maŋ
2	jau	haŋ[15]	men	ʈhaŋ					we	muŋ	ʈjaŋ	li	**fa**
3	lau	njin	soŋ	ŋai					ʈai	tjiu	to		
4	ka	tjiu	kɔp	ɕo					to	ti	**ʈha**		
5	so	ʈha	so	loi					hloi	soŋ	ʔit		
6	jit	so	kɐi	toŋ					jin	mɐi	loŋ	nai	**ha**
7	ʈhaŋ	lai	ʈhaŋ	ja					pau	ɕau	lau	poi	kaŋ
8	ta	hæn	ko	woi					poi	sən	**nja**		

Figure 6: The first half of the lyrics to *Lao nyin mang* (‘Old people happy’).

According to Yang & Zheng’s (1986: 252) general analysis[16] of Kam poetics, ending rhymes only occur between words with the linguistic tones 1, 2, 7 and 8; cross-line internal rhymes only occur between words with tones 3, 4, 5, 6, 9 and 10; and internal rhymes within the one line can occur between words with any tones. Wu & Zhang (1991) also note a correspondence between tonal categories and word rhymes in Kam songs. However, until there is a complete analysis of the tonal system of the second lect of the Southern Kam language it is difficult to confirm whether these analyses also hold for big song lyrics and other song genres specific to that region.

The process of Kam song transmission

The transmission of most genres of songs within the Kam village context usually involves a two-stage process: an initial stage of chanting song lyrics, followed by a stage of rendering them in song.

Kam songs as *yaoyun* 腰韵 (‘waist rhymes’), *gouyun* 勾韵 (‘hook rhymes’) and/or *neiyun* 内韵 (‘internal rhymes’).#

15 This syllable may be a coincidental rhyme, rather than strictly part of the rhyming scheme.

16 Yang & Zheng (1986: 250) state that their analysis is generally applicable for all Kam songs “besides some ancient ritual songs and *beeba* songs” (translation mine).

Stage 1: 'Getting the bones of the song' – Chanting the song lyrics

The first stage in this process is referred to as *ao lak ga*, an expression that literally translates as 'getting the bones of the song'. In this process the song expert chants one line of text and waits at the end of the line for the learners to chant the same line. The teacher then goes on to the next line, waiting for the learners to repeat it after her/him and continuing the same process for remaining lines. The teacher of the song usually goes through the entire song numerous times chanting it in this way, and learners only attempt to sing the song once they can recite the lyrics without the teacher's assistance. On occasions when the song has already been learnt and the singers want to revise the lyrics, they follow basically the same process but describe the activity as *dok ga* ('revising the song').[17] Often, *dok ga* involves all the singers chanting the lyrics together from start to finish.

Regardless of whether the activity is 'getting the bones of the song' or 'revising the song', in Sheeam one distinctive rhythm and a somewhat flexible sequence of non-tonally-fixed pitch variation is used for the chanting of all songs. As I have outlined elsewhere (see Ingram 2007: 85f.), the rhythm used to chant song lyrics seems to emphasize certain features of the rhyming structure of the lyrics, rather than emphasizing the rhythm used when the lyrics are sung. Figure 7 shows one chanted version of the first two lines of *Lao nyin mang*. Figure 8 shows a sung rendition of the same lyrics (a transcription of the entire first section is included in Appendix 2). The only obvious similarity between the chanted and sung versions is the lyrics. The rhythm used in the chanted version bears only minor, incidental similarity to the rhythm of the sung version, and the length of time required to chant each line is different from the length of time needed to sing each line. While there is some approximate correspondence between the pitch contours of these two examples, such correspondence is most probably incidental as it does not continue throughout the song.

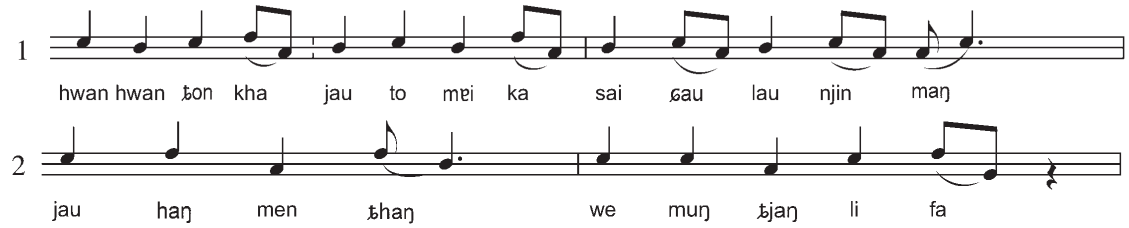


Figure 7: The tonal contour of the first two lines of *Lao nyin mang* ('Old people happy') as chanted by Sheeam song experts Sa Yuu-jin and Sa Yuun-yong (Sheeam, 18 May 2008). The chanted speed fluctuated between 180 and 200 crotchet beats per minute. Line 1 was chanted in 4.5 seconds; line 2 was chanted in 4 seconds. As the precise pitch is not fixed and is not important to the study here, the pitches indicated on this three-line stave represent only relative pitch contours (five different relative pitches are indicated, analogous to Chao Yuanren's system for notating the tonal system in Chinese).

17 Singers particularly use this method when revising lengthier songs that are tiring to sing in full.

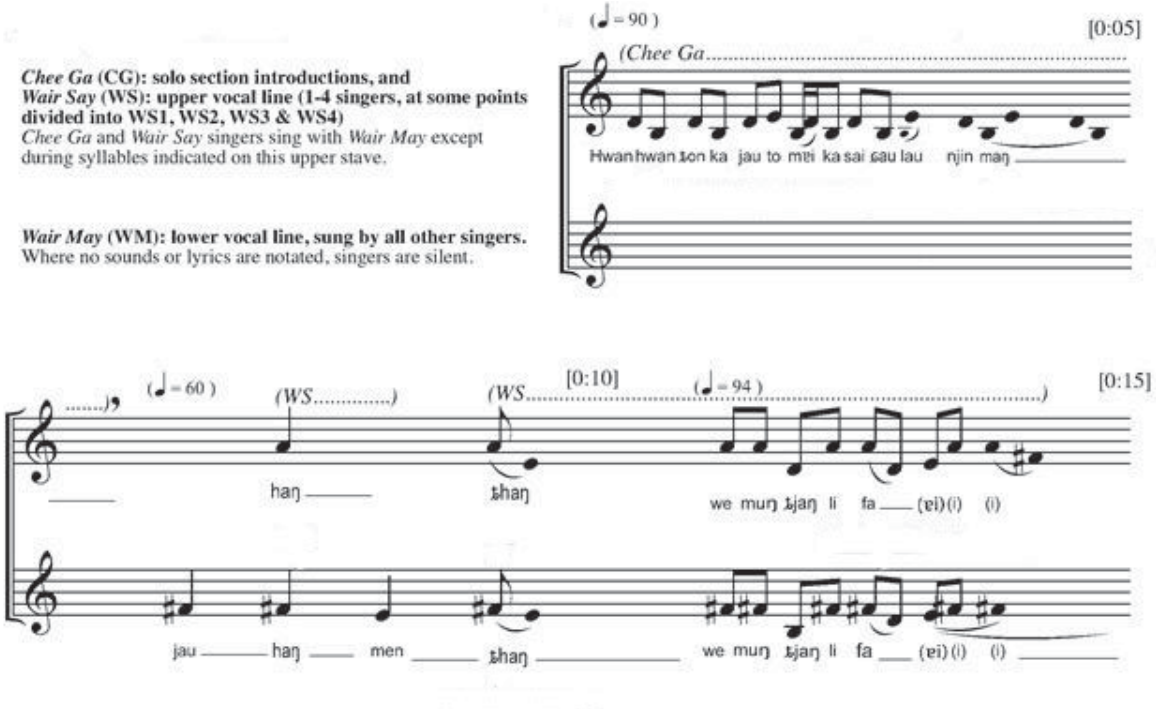


Figure 8: A transcription of a sung version of the same lines of *Lao nyin mang* appearing in Figure 7. This performance was given by a group of ten women of the Lak Jang clan resident in and around the Jooim Bare section of the Sheeam village called Jai Lao (recorded in the main *dare low* in Jai Lao on 8 February 2008). The *chee ga* (solo introduction to song sections) and the *wair say* (upper vocal line) are transcribed on the upper stave. In this performance, the length of time used to sing the second line of lyrics (appearing here on the second pair of staves) was more than twice the time required to chant it. In this transcription, each full musical system (a pair of horizontal staves) represents ten seconds of music, as marked above the staves in square brackets in the form [minutes:seconds]. Note heads are placed in a linear arrangement relative to their position in time, and note heads without stems are used in unmetered sections (see also Ingram 2012 and Appendix 2).

This Kam use of chanted lyrics in the transmission of a folk song genre may be somewhat unusual, particularly within a Chinese context. For example, in her study of Jiangsu *shan'ge* 山歌, Antoinet Schimmelpenninck (1997: 404) states: "[…] spoken versions of [these] song texts are artificial constructs. Singers do not learn or acquire the lyrics in spoken form, but usually *as* music". Due to the distinctive process of Kam musical transmission, it would appear that Kam singing traditions thus differ in a significant way from the important *shan'ge* folk song tradition of one Chinese region.[18]

18 The "speech–song continuum" is well recognized (see List 1963; Wadley 1991: 201; Briggs 1993: 934; Feld 1990: 252–254; Feld et al. 2004: 332; Seeger 2004 [1987]). Where the chanting of big song lyrics would be placed in relation to two extremes of "speech" and "song" is an issue worthy of further consideration.

Stage 2: 'Teaching/singing the song'

The second stage of song transmission involves rendering the chanted lyrics in a way that Kam people recognize as being a "song". This is achieved by singing the lyrics with an appropriate *sor*. I sometimes heard singers refer to this process as *gay-uu* ('teach[ing]') the song, whilst on other occasions it was simply known as learning how to *dor* ('sing, do, carry out') the song.[19] Occasionally – particularly when a number of singers were learning a choral song to perform together as a group, or less experienced singers were learning a song – in this second stage the singers learnt to create the sung version by once again copying each line of lyrics after it had been sung by the teacher.

However, in most instances the learners' prior familiarity with the *sor* of songs in their local repertoire meant that they had very little difficulty in all singing the song together once they had all memorized the song lyrics. This suggests that it was typical for singers to have unconsciously absorbed the principles concerning the STM relationship. Sometimes, learners were so confident of their ability to produce a sung version of chanted song lyrics that they did not even sing through the song at all. This was often the case when experienced singers learnt extra songs in order to lead a group in call-and-reiteration singing at a forthcoming engagement or wedding, or in celebrations for a new house or a new baby. Actual singing of newly learnt song lyrics was also unnecessary when the song being learnt was a solo song (such as *ga nya* or *ga bai-jin*). In those instances, the learner was expected to have her/his own individual manner of *yuk* or 'rendering' the *sor*, as described by Sheeam song expert Sa Yuu-jin:

> *Gak moong gak shyop yuk e-no-yang shyuu yuk e-jey-yang* ('[For] each person will *yuk* [i.e. render the *sor*] in the way that they prefer') (Sa Yuu-jin, personal communication, March 2011)

Finally, when singers were involved in learning songs such as big songs that have two simultaneous vocal lines that are performed by different members of the one singing group, the process of teaching how to sing the song was itself divided into two stages: a stage when all singers learnt to sing the lower vocal line, and a subsequent stage when the upper vocal line was taught and added.

Discussion

I focus this preliminary analysis upon two Kam songs with different performance formats – *Lao nyin mang* ('Old people happy') and *Deeuu kun*

19 The Kam word *dor* also has many more different meanings, depending on the context in which it is used (see Ingram 2010).

nai ghai ('This road is long')[20] The first song, *Lao nyin mang*, is a big song. It therefore involves the simultaneous performance of two different vocal lines, and is performed by a well-rehearsed group of singers. The second song, *Deeuu kun nai ghai*, is always performed by a solo singer. While the lyrics to this song have only one chanted form, when they are rendered in song they have two different sung forms as they can be performed in two different genres: *ga bai-jin* and *ga nya*. The two different *sor* which can be used to render the one set of lyrics to this second song provide an interesting and useful point of comparison in analyzing the STM relationship.

Chanted song lyrics

In my preliminary analysis of chanted song lyrics I examine three aspects of the chanted lyrics: the degree of similarity between the pitch contours of the spoken and chanted forms of the lyrics; the degree of similarity in the overall conception of spoken and chanted forms of the lyrics; and whether or not linguistic tone might be linked to the rhyming patterns in the lyrics.

1. Detailed analysis of pitch contour in chanted song lyrics

One of the most important questions regarding chanted forms of song lyrics (and thus the process of musical transmission in Kam singing traditions) is the degree to which the chanted pitch contours correspond with the linguistic tones used when the same words are spoken in everyday speech. However, because there is no comprehensive linguistic data regarding the second lect of Southern Kam it is difficult to present the spoken and the chanted forms of the song lyrics for comparison. Furthermore, during my fieldwork I also discovered that singers found it virtually impossible to speak the words of song lyrics as they would in daily speech. Their difficulties in simply speaking the lyrics were due to the fact that the structural form of the lyrics often differed from expressive forms used in everyday speech, and because of the singers' great familiarity with reciting the lyrics in a chanted form.

Despite these difficulties, I developed an initial methodology through which some aspects of the relationship between tones of spoken and chanted words can be inferred in at least a preliminary sense. In this methodology I refer to the meanings of words in the lyrics to ascertain the multiple appearances of the one word in the same set of lyrics, and then compare the pitch contours used in chanting the different recurrences of the same word. Additionally, for

20 The lyrics to this song can be translated as follows: 'This road is long / I always call on you to make it shorter / These mountain valleys are deep / I always call on you to make them level / If the mountains cannot be level / I will ask you to place stones to fill the gaps / If the road is difficult to traverse / I will ask you to build a bridge / Build a wooden bridge / Later it will rot / Build a stone bridge / Afterwards it will be standing forever.'

many common words in the Sheeam variant, I am also able to compare the chanted pitch contours of these words with my own understanding of the most common appropriate linguistic tone of each word as developed through my experience with and fluency in the Kam language.

Whilst the pitch contours in both speaking and chanting the same word may vary when the word is used in different contexts – for example, there may be complex tonal variations as a result of factors including intonational pitch change, tone sandhi (influences of the tones of neighbouring words or syllables), the overall intonation of an utterance, and the functional load of tones – I suggest that an analysis such as that proposed here nevertheless provides a useful starting point for discussing the correspondence between spoken and chanted pitch contours in greater detail. In Figure 9, I compare four of the words that recur within the lyrics to each of the two songs. As is evident from the final line in each table, I believe that the words in these songs that recur have different lengths of vowels when spoken in daily speech. Research beyond this preliminary stage would ideally involve a wider selection of words with long vowels, since it can be assumed that long vowels allow greater scope for rendering the tonal contour.

1. Lao nyin mang:

Kam words/ expressions:	*ka* - /ka/ (‘song’)	*ting* - /thiŋ / (‘listen, hear’)	*lau nyin* - / lau ŋin/ (‘old people’)	*nyum nai* - / ŋɐm nai/ (‘tonight’)
First occurrence	**5→2**	**5→3**	**3, 4→2**	**4→2, 2**
	(end A, L1)	(end B, L13)	(middle B, L1)	(start A, L11)
Second occurrence	**4→3**	**3→4**	**2, 3**	**4, 2**
	(start A, L4)	(start B, L16)	(start A, L3)	(start A, L16)
Third occurrence	**4**			
	(middle B, L11)			
Fourth occurrence	**5→3**			
	(end A, L13)			
Fifth occurrence	**4→1**			
	(end B, L16)			
My under-standing of the tones usually used in the Sheeam vari-ant	5→1 (guttural, with short vowel)	5→3 (long vowel)	3, 4→2 (two long vowels)	4→2, 2 (short vowel, then long vowel)

2. Deeuu kun nai ghai:

Kam words/ expressions:	*kun* - / khwən / (‘road’)	*nya* - /ɲa/ (‘you’)	*jeeuu* - / ʨiu / (‘bridge’)	*deen* - /tin / (‘stone, rock’)
First occurrence	**3**	**4→1**	**4→2**	**4→2**
	(middle A, L1)	(middle B, L1)	(middle A, L5)	(middle B, L3)
Second occurrence	**4**	**4→1**	**4**	**5→1**
	(start A, L4)	(middle B, L2)	(middle A, L6)	(end A, L6)
Third occurrence		**4→1**		
		(middle B, L4)		
My under-standing of the tones usually used in the Sheeam vari-ant	4→1 (short vowel)	3→4→1 (long vowel)	4→2 (long vowel)	5→1 (short vowel)

Figure 9: A comparison of the pitch contours (given numerically, in bold) used for chanting four Kam words or expressions at the different times they recur in the lyrics to *Lao nyin mang* (see Appendix 1 for a full transcription of the chanted lyrics to this song) and *Deeuu kun nai ghai*. The different occurrences of the one word or expression are placed in the same column, and my understandings of the tones usually used for each word or expression in the Sheeam variant appear at the bottom of each column. The precise location of each occurrence within the song lyrics is given in brackets below each numerically represented pitch contour (for example, “end A, L1” means that the word or expression occurs at the end of the “A”-section in the first line of lyrics). Words or expressions with a chanted pitch contour that differs significantly from the contour I suggest is usually used in daily speech are shaded dark grey. Words or expressions that form part of the rhyming pattern of the song lyrics are shaded light grey.

In the examples from *Lao nyin mang*, in one instance *ka* is chanted without a tonal descent (although it does use the high pitch characteristic for the beginning of the pitch contour of this word), whilst in all other contexts (including daily speech) the pitch usually descends on this word. Furthermore, amongst all chanted forms of pitch descent for *ka*, the shape of the contour varies, but none have a descending contour as marked as that usually used for the word in daily speech. In the case of the word *ting*, the two occurrences present quite different pitch contours, and only one corresponds with the pitch contour usually used in daily speech. In a general sense the two chanted forms of *lau nyin* and *nyum nai* are quite similar, but in both cases the precise details of tonal movement within the expressions differ. While there is no immediately obvious pattern regarding the differences amongst occurrences of each word or expression, it is interesting that for all cases where the word or expression is involved in the rhyming pattern there

seems to be close correspondence between the chanted pitch contour and the pitch contour most commonly used for it in daily speech.

In the examples from *Deeuu kun nai ghai*, neither instance of the chanted word *kun* involves the tonal fall clearly obvious in daily speech, though both chanted words begin at a pitch height similar to the height used at the start of the fall used in speaking the word. The three occurrences of the word *nya* all approximate the characteristic fall in this word when it is spoken, although the slight initial rise that is very important in the spoken form for distinguishing 'you' (*nya* with the tonal contour 341) from 'river' (*nya* with the tonal contour 51) appears to be absent in the chanting of this word. One occurrence for *jeeuu* lacks the characteristic tonal fall, though it does begin from the pitch used at the start of the word in daily speech. The two chanted forms of *deen* exactly or closely mirror the pitches used in speech.

2. Detailed analysis of the overall conception of spoken and chanted forms of lyrics

The analysis below examines each pair of lines in turn and compares three features: the number of syllables, the number of chanted pulses and the chanted length of time. In ordinary speech, the number of syllables in a phrase would be expected to directly relate to the time taken to speak the phrase. However, this preliminary analysis demonstrates that such a principle does not hold in chanted forms of the lyrics.

As noted above, the rhythm for a chanted rendition of a set of lyrics differs completely from the rhythm used in its sung form, and the time taken to chant a line can similarly differ. The following table, Figure 10, presents a comparison of the three features (syllables, pulses and actual length in time) for each pair of lines of lyrics in the song *Lao nyin mang*.

Pair of lines	**Number of syllables in each (odd/even lines)**		**Number of pulses in each in this rendition (odd/even lines)**		**Actual duration of each line in this rendition (secs.) (odd/even lines)**	
1 and 2	8	9	14	11	4.5	4.0
3 and 4	7	7	9	8	3.5	2.5
5 and 6	7	9	8	11	2.5	2.5
7 and 8	9	7	11	9	3.5	2.5
9 and 10	9	9	11	11	3.5	3.0

11 and 12	9	9	10	11	3.5	2.5
13 and 14	7	9	8	11	3.0	2.5
15 and 16	11	11	13	12	n/a	3.0

Figure 10: Table showing syllables, pulses and line durations in pairs of lines of the same chanted version of *Lao nyin mang* analyzed above in Figure 9. "Pulse" is used to refer to the sense of regular beat (represented by crotchet pulses in the transcriptions of the chanted form of the song given in Figure 7 and Appendix 1). Pairs of lines with fewer syllables in the first line of the pair are shaded light grey, while those with more syllables in the first line of the pair are shaded dark grey (pairs of lines with equal numbers of syllables in each are not shaded); n/a indicates a line where the singers momentarily forgot the lyrics and took extra time to remember them.

The data presented here shows that the length of time taken to chant the second line of lyrics in a pair (that is, an even-numbered line of lyrics) is usually less than the time required for the first line of a pair, regardless of the actual number of syllables or pulses in each line. The only exception in this example is lines 5 and 6, where the same amount of time was used for each line.

3. Detailed analysis of the link between linguistic tone and rhyming patterns

As mentioned previously, several researchers have identified a link between linguistic tone and rhyming patterns in the Kam songs of other regions. This issue is important for understanding the STM relationship: it may explain the choice of words with certain tonal contours at certain points in the lyrics, and it may influence the pitch contours of the melody. Without comprehensive linguistic data the tonal patterns within the spoken version of the song lyrics cannot be presented and how they might link with rhyming patterns cannot be discussed. However, it is possible to identify patterns of pitch contours related to rhyming patterns within the chanted version of the lyrics.

In this particular chanted rendition of the lyrics to *Lao nyin mang*, all but one of the eight ending rhymes involve a pitch descent (four involve 5→1, two involve 4→1, one involves 3→1), suggesting that a similarity in chanted pitch contours may be important for ending rhymes. The pitch contours for chanting cross-line internal rhymes are more varied, as are those for the few internal-line internal rhymes. As shown in the table in Figure 11, in four of the six internal-line internal rhymes a closely similar pitch contour was used. However, amongst the ten instances of cross-line internal rhymes, less than half displayed similarity.

Pairs of lines	Pitch contours used for chanting cross-line internal rhymes				Pitch contours used for chanting internal-line internal rhymes (always in odd-numbered line)		
1 and 2	2→4	5→3	2		5→2	5→2	
3 and 4	5	4→5	4		5	4	
5 and 6	3→5	4		(and also another rhyming pattern with pitch contours 5→3, 2)	2	4	
7 and 8	2	1			4	3	3→4
9 and 10	1	1	1		5	4	5
11 and 12	5→2	1	2				
13 and 14	5→3	3	1				
15 and 16	5	2→4		(and also another rhyming pattern with pitch contours 1, 1)	3→4→3	4	5

Figure 11: Patterns of linguistic tones used in internal rhymes in the same chanted version of *Lao nyin mang*. The rhymes which have identical or closely similar pitch contours are shaded.

The chanted rendition of *Deeuu kun nai ghai* presents an interesting comparison. As evident from the transcription of the lyrics to this song provided in Appendix 3, the lyrics to the song have no rhyming patterns that continue in such a structured way throughout the entire song. Instead, the last syllable of the first and final lines rhymes includes the sound "in," but the tonal contours used for chanting these two syllables are entirely different. Otherwise, there are occasional rhymes between the last syllable of a line and a syllable within the next line, approximating the cross-line internal rhymes described above (although without the strict regulations concerning the precise pairing arrangement of cross-line internal rhymes as beginning only on odd lines).

Preliminary conclusions regarding chanted lyrics

This preliminary analysis indicates that while the same word can be chanted with different pitch contours in different contexts, a large degree of deviance from pitch contours used in daily speech appears uncommon. Moreover, words or expressions that are involved in the rhyming pattern seem to be chanted with a contour particularly close to the spoken form.

Preliminary research suggests that singers perceive even-numbered lines quite differently from odd-numbered lines when chanting lines of lyrics, and that there is no clear relationship between the number of syllables in a line and the number of pulses felt when chanting the line. This study identifies conceptualization of spoken and chanted forms of lyrics as an important area for future research. Such further research will help to determine whether there are also similar differences in the ways that singers conceptualize the sung form of each line; whether the way the STM relationship functions or should be rendered in odd-numbered lines of sung forms might be different from the ways it functions or is rendered in even-numbered lines; whether felt pulses are an important consideration when chanting lyrics; and whether this situation is indicative of other organizational principles not immediately evident that may have a bearing upon the STM relationship within sung forms.

This analysis suggests that there may be specific pitch-related requirements for the chanted pitches of ending rhymes, and perhaps also for internal-line internal rhymes, but probably not in the case of cross-line internal rhymes. If present, these requirements might also influence the choice of words in the song lyrics. However, the situation regarding the use of descending pitch contours for ending rhymes requires some qualification, as the Sheeam variant of the Kam language is well-known for its preponderance of falling tones, so the regularity of falling tonal contour patterns is somewhat to be expected. I suggest that the degree to which rhymes in song lyrics are based on similar tonal patterns still requires considerable attention. My initial view in this regard is not only based on the results of the analysis presented here. During my fieldwork, I found that confusion in the interpretation of the lyrics was often because singers disputed the appropriate linguistic tone (and thus meaning) of a word. While such confusion concerning analysis of song lyrics within tonal languages is expected, the degree to which it occurred during my research cast doubt on the existence of a straightforward, regular underlying tonal pattern to the rhyming scheme, which would be expected to somewhat reduce such confusion.

Comparing sung and chanted song lyrics

Comparison of sung and chanted song lyrics is the second crucial aspect in understanding how the process of musical transmission influences the STM relationship in Kam song. The variety in performance formats used in Kam singing is an important consideration for comparing sung and chanted song lyrics, since different performance formats place different requirements upon singers regarding the way they might render the *sor*.

Detailed analysis of the comparison between chanted and sung lyrics

As previously mentioned, clear differences between chanted and sung lyrics include the rhythm used and the length of time required to chant or sing each line. While I have noted some similarities between the two versions – such as a general correspondence in the pitch contours for the first two lines of lyrics – such correspondence does not continue throughout. While in the first two lines of lyrics some syllables involved in the rhyming structure seem to be sung and chanted with relatively similar pitch contours, in later sections of the song the syllables involved in the most prominent rhymes – ending rhymes – are some of the syllables that exhibit the greatest degree of difference between chanted and sung forms. Overall, from this particular example of big song singing, it is difficult to identify any very strong or obvious correspondence between pitch contours used in sung and chanted versions of the song.

The data derived from analyzing one set of chanted lyrics for the song *Deeuu kun nai ghai* and two different sung versions of the same lyrics (one version as *ga nya* and one version as *ga bai-jin*) presents a somewhat different picture. Unlike the choral big song *Lao nyin mang* discussed above, both the *ga nya* and the *ga bai-jin* versions of *Deeuu kun nai ghai* are sung by just one singer, allowing substantial flexibility in the rhythm and melodic contours of the *sor* used in each case. The pitch range of both sung versions is mainly within a five-note range, therefore the chanted version and the basic structure of the two sung versions can be mapped onto the same chart for comparison. The pitch contours of the various versions of the two sections of the first line of the song are graphically represented in Figure 12. Here, for ease of comparison, each syllable is nominally allotted three time units along the horizontal axis. The pitch contour of the chanted form is marked in blue, while red and green are used to indicate the pitch contours of the *ga nya* and *ga bai-jin* versions respectively. A pitch scale of 1-7, based around the 1-5 tonal range used above, allows accommodation of occasional extensions beyond the 1-5 range in the two sung versions of the song.

Section 1 of Line 1

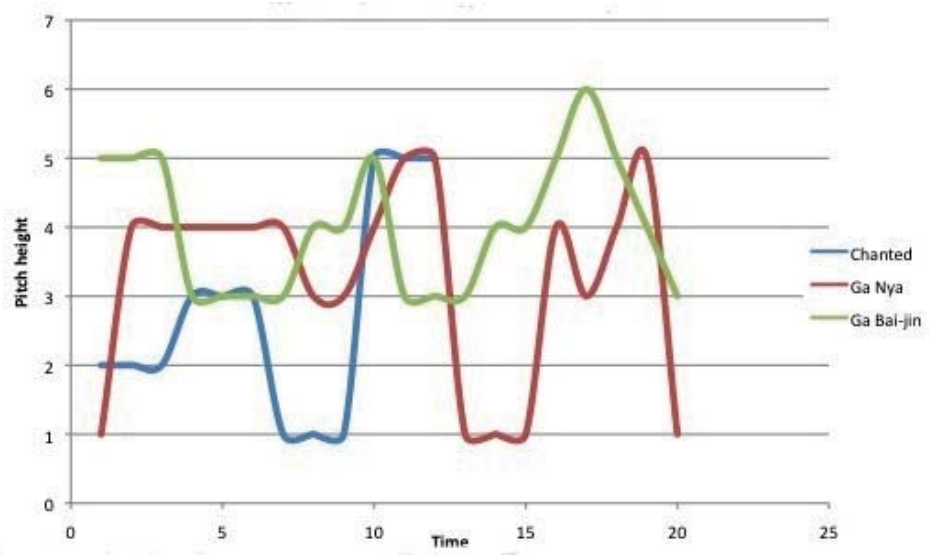


Figure 12a/b: A comparison of three different versions of the two sections comprising the first line of lyrics to the song *Deeuu kun nai ghai* (transcribed from a recorded performance by Sa Yuu-jin in Sheeam, 7 March 2011). See Appendix 3 for a lengthier mapped comparison.

Section 2 of Line 1

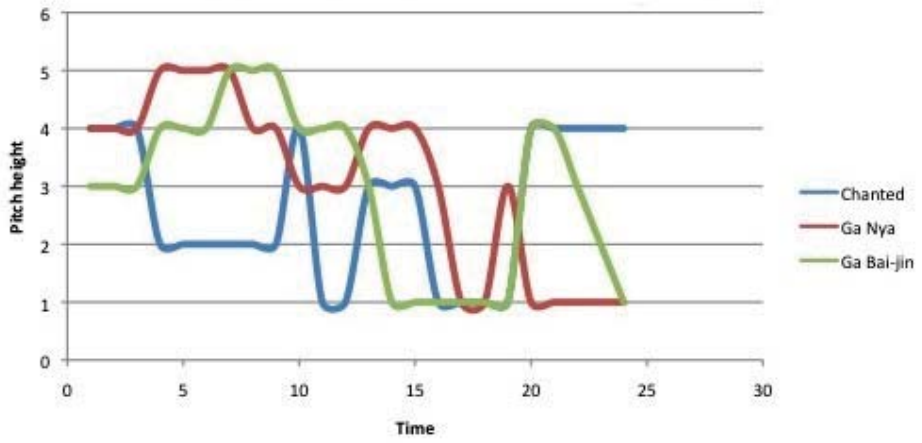


Figure 12b

This method for comparing the two different *sor* used to sing the same line of lyrics with the chanted form of the lyrics aids identification of the features of the melodic line that might be linked to the lyrics and the features of the melodic line that might be important in distinguishing the particular *sor*. As evident from Figure 12a, the pitch contours of all three versions of the song rise and meet in the last syllable of the actual lyrics (over time units 10-12). They then differ over the subsequent syllables – all of which are vocables and thus not part of the chanted version of the song. In figure 12 b, showing the second section of line 1, the three pitch contours all follow a generally descending pattern that is followed by a final short rise.

In the second line of lyrics (see Appendix 3), the pitch contours of all three versions of the song have a general rise from a relatively low starting pitch at

the beginning of the line. The three pitch contours only deviate widely at the end of the line. In the second section of the second line two general descents are apparent, and the pitch contours for all three versions of the song finally exhibit a brief rise and fall to end on the lowest pitch.

Preliminary conclusions on the comparison between chanted and sung lyrics

In the case of the big song examined here it is difficult to identify any very strong or obvious correspondence between pitch contours used in sung and chanted versions of the song. Further research would be required to prove whether the occasional similarities in the contours of the first two lines of these two versions are incidental or are an occasional feature of songs in this genre.

However, a comparison of the pitch contours used for the chanted lyrics to *Deeuu kun nai ngai* with the contours used when the same lyrics are sung both as *ga bai-jin* and as *ga nya* presents a different case. While it is clear that the pitch contours of the chanted version are not precisely replicated in either of the sung versions, in this example certain general similarities amongst the versions are apparent. These results suggest that in solo song performances it may be overall pitch contours, rather than the pitch contours used for particular syllables, that present the greatest similarities amongst various versions of a set of song lyrics. It is possible that the chanted version of the lyrics provides a kind of blueprint for adapting spoken pitch contours into a sung form, supplying the singer with information about which spoken pitch contours are most important and should be retained in the sung renditions.

Conclusion and further investigation

This preliminary analysis demonstrates that the pitch contours used for both chanted and sung forms of song lyrics in Kam song traditions do not precisely replicate the pitch contours used for the same words in daily speech. However, the present study also shows that there are indeed various kinds of resemblances amongst spoken, chanted and sung forms of song lyrics. It confirms that Kam singers are not solely relying on their knowledge of the linguistic tones of the words used in daily speech when they render the same words in sung form, and that the chanted form of song lyrics – and hence the process of musical transmission – is linked to the pitch contours of spoken and sung lyrics in a variety of ways. Clearly, investigation of the STM relationship in Kam song traditions is a complex endeavour that demands recognition of the role of the process of musical transmission. This analytical endeavour is made even more challenging by the two significant factors mentioned at the beginning of the paper: insufficient specific linguistic data on the second lect of Southern Kam, and inadequate prior documentation of the process of Kam song transmission.

However, the present study outlines a range of methodologies that can be used to at least partially overcome these challenges. It demonstrates that the data generated through utilizing these methodologies is very valuable in analyzing the STM relationship in Kam song traditions, and particularly in analyzing how the STM relationship is influenced by the process of musical transmission.

On the basis of this preliminary analysis, it can be concluded that while the process of musical transmission is influential upon the STM relationship in Kam singing traditions, its influence varies across different traditions. In the case of big song, where the musical lines are sung by a group of pre-rehearsed singers, the pitch contours of the spoken and chanted forms of the song lyrics appear to resemble each other much more closely than those of the chanted and sung forms. Further research is required to determine if this means that the main role of the process of musical transmission (the chanted form) is to help singers interpret and hence recall the lyrics. Further research will also reveal how the chanted form of song lyrics may have a bearing upon the different ways in which the one *sor* may be rendered, and how other important features of the lyrical structure and transmission – such as the involvement of linguistic tone within the rhyming structure and the different conceptualizations of different lines of lyrics (as evident in the chanted form of the lyrics) – might similarly be related to the pitch contours when the lyrics are rendered in sung form.

In the case of a single set of lyrics being performed with different melodies in two different song genres, a very different picture emerges regarding the relationship between chanted and sung lyrics. In such cases it appears that overall pitch contours of each section of a line of lyrics have a general resemblance, and that the chanted form of the lyrics may have an important role in helping singers to determine the overall tonal constraints as they create their own individual rendering of the *sor*. Similarly, further research will help to explicitly outline how other important features of the lyrical structure and transmission – such as the involvement of linguistic tone within the rhyming structure and the different conceptualizations of different lines of lyrics (as evident in the chanted form of the lyrics) – might be related to the pitch contours when the lyrics are rendered in sung form.

Further research that draws on a wider range of recorded material will help to develop these preliminary conclusions, and to further explore the many potential relationships between speech tone and melody concerning Kam song traditions. Analysis of multiple versions of the same set of song lyrics as sung in different song traditions, and as performed by different singers, will be particularly helpful in further inquiry. Most importantly, any results of this analysis will need to be tested through being used to inform performance when singing together with Kam singers.

Please visit http://www.phonogrammarchiv.at/7DE709JKSAD
for additional appendices.

References

Bourdieu, Pierre. 1977 [1972]. *Outline of a Theory of Practice*. Translated by R. Nice. Cambridge etc.: Cambridge University Press.

Briggs, Charles L. 1993. "Personal Sentiments and Polyphonic Voices in Warao Women's Ritual Wailing: Music and Poetics in a Critical and Collective Discourse". *American Anthropologist* 95 (4): 929–957.

Diller, Anthony V. N., Edmondson, Jerold A. & Luo, Yongxian (eds.). 2008. *The Tai-Kadai Languages*. Oxford & New York: Routledge.

Edmondson, Jerold A. & David B. Solnit (eds.). 1988. *Comparative Kadai: Linguistic Studies Beyond Tai*. (Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics, 86). Dallas: Summer Institute of Linguistics & The University of Texas at Arlington.

Feld, Steven. 1990. "Wept Thoughts: The Voicing of Kaluli Memories". *Oral Tradition* 5 (2–3): 241–266.

Feld, Steven et al. 2004. "Vocal Anthropology: From the Music of Language to the Language of Song". In: Duranti, A. (ed.). *A Companion to Linguistic Anthropology*. Malden, MA etc.: Blackwell Publishing, 321–346.

Gedney, William J. (ed. by Thomas John Hudak). 1997. *William J. Gedney's Thai and Indic Literary Studies*. (Michigan Papers on South and Southeast Asia). Ann Arbor, Michigan: Center for Southeast Asian Studies, University of Michigan.

Gerner, Matthias. 2009. "Instruments as Verb Classifiers in Kam (Dong)". *Linguistics* 47 (3): 697–742.

Holm, David. 2003. *Killing a Buffalo for the Ancestors. A Zhuang Cosmological Text from Southwest China*. Illinois: Southeast Asia Publications.

Huang, Yongsha 黄勇刹. 1987. "Zhuangzu minjian shilu chutan" 壮族民间诗律初探 [An Initial Exploration of Zhuang Folk Poetics]. In: Duan, Baolin 段宝林 & Guo, Wei 过伟 (eds.). *Minjian shilu* 民间诗律 [Folk Poetics]. Beijing: Beijing daxue chubanshe [Beijing University Press], 135–155.

Hudak, Thomas John. 2008. "Tai Aesthetics". In: Diller, Anthony V.N., Edmondson, Jerold A. & Luo, Yongxian (eds.). *The Tai-Kadai Languages*. Oxford & New York: Routledge, 404–414.

Ingram, Catherine. 2007. "'If You Don't Sing, Friends Will Say You are Proud': How and Why Kam People Learn to Sing Kam Big Song". *Context: A Journal of Music Research* 32: 85–104.

– – –. 2010. Hwun hwun jon ka *(Listen) … Kam Villagers Singing Kam Big Song in Early Twenty-First-Century China*. PhD thesis, University of Melbourne.

– – –. 2011. "Echoing the Environment in Kam Big Song". *Asian Studies Review* 35 (December): 439–455.

– – –. 2012. "*Ee, mang gay dor ga ey* (Hey, why don't you sing)? Imagining the Future for Kam Big Song". In: Howard, K. (ed.). *Music as Intangible Cultural Heritage: Policy, Ideology and Practice in the Preservation of East Asian Traditions*. Farnham, UK & Burlington, VT, USA: Ashgate, 55–75.

Ingram, Catherine & Wu, Zhicheng 吴志成 (forthcoming). "*Beeba* (Kam Pipa)". In: Libin, L. (ed.). *The New Grove Dictionary of Musical Instruments*. Oxford: Oxford University Press.

Liang, Min 梁敏 & Zhang, Junru 张均如. 1996. *Dongtai yu gailun* 侗台语概论 [An Introduction to the Kam-Tai Languages]. Beijing: Zhongguo shehui kexue chubanshe [China Social Sciences Press].

List, George. 1963. "The Boundaries of Speech and Song". *Ethnomusicology* 7: 1–16.

Long, Yaohong 龙耀宏 & Zheng, Guoqiao 郑国桥. 1998. *The Dong Language in Guizhou Province, China*. Translated by D.N. Geary. (Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics, 126). Dallas: Summer Institute of Linguistics & The University of Texas at Arlington.

Lundström, Håkan. 2010. *I Will Send My Song: Kammu Vocal Genres in the Singing of Kam Raw*. Copenhagen: NIAS Press.

Matthews, Stephen. 2006. “Cantonese Grammar in Areal Perspective”. In: Aikhenvald, A.Y. & R.M.W. Dixon (eds.). *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 220–236.

Miller, Terry E. 1985. *Traditional Music of the Lao: Kaen Playing and Mawlum Singing in Northeast Thailand*. Westport, CT: Greenwood Press.

Mittler, Barbara. 2001. Book Review: Antoinet Schimmelpenninck. *Chinese Folk Songs and Folk Singers. Shan’ge Traditions in Southern Jiangsu*. *China Information* 15 (1): 314–316.

Pan, Yongrong 潘永荣. 2005. “Liping xian Yongcong xiang Jiulong cun Dongyu diaocha baogao” 黎平县永从乡九龙村侗语调查报告 [A Research Report on the Kam Language of Jiulong Village, Yongcong Township, Liping County]. In: *Guizhou minzu diaocha juan ershiyi 2003* 贵州民族调查卷二十一2003 [The Twenty-first Volume of Research into Guizhou’s Ethnic Groups, 2003]. Guiyang: Guizhou sheng minzu yanjiusuo [Nationalities Research Centre, Guizhou Province] & Guizhou sheng minzu yanjiuhui [Nationalities Research Council, Guizhou Province], 486–503.

Sagart, Laurent. 2004. “The Higher Phylogeny of Austronesian and the Position of Tai-Kadai”. *Oceanic Linguistics* 43 (2): 411–444.

Schimmelpenninck, Antoinet. 1997. *Chinese Folk Songs and Folk Singers: Shan’ge Traditions in Southern Jiangsu*. Leiden: CHIME Foundation.

Seeger, Anthony. 2004 [1987]. *Why Suyá Sing: A Musical Anthropology of an Amazonian People*. Urbana & Chicago: University of Illinois Press.

Thrasher, Alan R. 1990. *La-li-luo Dance-songs of the Chuxiong Yi, Yunnan Province, China*. Danbury CT: World Music Press.

Wadley, Susan S. 1991. “Why Does Ram Swarup Sing? Song and Speech in the North Indian Epic Dholā”. In: Appadurai, A., F.J. Korom & M.A. Mills (eds.). *Gender, Genre and Power in South Asian Expressive Traditions*. Harrisburg: University of Pennsylvania Press, 201–223.

Wu, Hao 吴浩 & Zhang, Zezhong 张泽忠. 1991. *Dongzu geyao yanjiu* 侗族歌谣研究 [Research into Kam Song]. Nanning: Guangxi renmin chubanshe [Guangxi People’s Press].

Yang, Quan 杨权. 1986 [1982]. “Lun Dongzu shige de yinlu tese” 论侗族诗歌的韵律特色 [On the Special Characteristics of the Rhyming Patterns in Kam Poems and Songs]. In: Wang, Shengxian 王胜先 (ed.). *Dongzu wenhua shiliao 1–10 juan* 侗族文化史料1－10卷 [Historical Materials in Kam Culture, Vols. 1–10]. Guiyang: Qiandongnan Miaozu Dongzu zizhizhou minwei minzu yanjiusuo [Nationalities Research Office of the Southeast Guizhou Miao & Kam Autonomous Prefecture Nationalities Commission], 168–172.

Yang, Quan 杨权 & Zheng, Guoqiao 郑国乔. 1986. “Dongzu shige yunlu” 侗族诗歌韵律 [Rhyming Schemes of Kam Songs and Poems]. In: Zhongyang minzu xueyuan shaoshu minzu wenxue yishu yanjiusuo wenxue yanjiushi 中央民族学院少数民族文学艺术研究所文学研究室 [Literature Research Office of the Nationalities Literature and Arts Research Centre at the Central Nationalities Institute] (ed.). *Shaoshu minzu shige gelu* 少数民族诗歌格律 [Tonal and Rhyming Patterns in the Poems and Songs of Minority Groups]. Beijing: Xizang renmin chubanshe [Tibetan People’s Press], 250–264.

Zhang, Yong 张勇 & Liu, Zhenguo 刘振国 (eds.). 1997. *Dongzu quyi yinyue* 侗族曲艺音乐 [Kam Narrative Art Music]. Guiyang: Guizhou minzu chubanshe [Guizhou Nationalities Press].

Zheng, Guoqiao 郑国乔 & Yang, Quan 杨权. 1988. “The Sounds of Rongjiang Kam”. In: Edmondson, Jerold A. & David B. Solnit (eds.), 43–58.

Speech and song: investigating the borderland

(Anastasia Karlsson, Håkan Lundström, Jan-Olof Svantesson & Siri Tuttle)

This article deals with language and music in two quite different languages and music cultures: that of the Kammu people in Laos and that of two Athabascan groups in Alaska. It has come out of a research project concerned with the study of vocal expressions in the borderland between speech and song, which spans over several language and music settings in societies where oral transmission of culture dominates. The aim of the project is to increase knowledge through collaboration between researchers with different approaches, to develop an interdisciplinary method for analysis of such expressions and to use this method in an intercultural study including several language and music settings. A long-term aim is also to play a part in the revitalization of such oral traditions and to contribute to their sustainability.

These two language settings have been chosen because they are quite different, as are the respective cultures. Our aim is not comparison of their vocal expressions, but rather an attempt to investigate the nature of the borderlands between speech and song in these two cultures. This choice does not represent a universalist claim about the relationship between speech and song; rather it sets up a research space in which questions about universality and variation may be investigated.

The study of the relation between speech and song in Kammu culture has so far focused on poetic metrics and on the musical consequences of the existence of lexical tones, whereas the influence of prosody and singing in non-tonal Kammu dialects are more recent areas. Though the Athabascan languages studied here do not have tone associated with every syllable, lexical tones do occur and their relation to music is a new area of study as is the relation of language and music in the structural constructions of the songs. This article is an attempt to further integrate linguistics and musicology and also a first step in the search for a cross-cultural approach to the understanding of practices for re-creation of spoken and/or sung statements.

Speech and song

In everyday situations the meaning of ‘speech’ and ‘song’ are quite commonly understood, or when focusing on the action: ‘speaking’ and ‘singing’. When these terms are not precise enough, other descriptions

are normally used, for example nouns like ‘declamation’, ‘recitation’, ‘incantation’, ‘chant’, or corresponding verbs like ‘declaiming’, ‘reciting’, ‘chanting’. Thus the English language has a number of terms to describe related phenomena, but they are not very precise. Attempts have been made to organize such concepts on a scale spanning from speech to song (List 1961, 1963).

Turning to music traditions in cultures where music is mainly orally transmitted, terms like ‘song’ or ‘singing’ are problematic. They are in themselves ethnocentric and therefore in many ways misleading. One reason for this is that in the study of orally transmitted singing the concept ‘song’ has been closely associated with the concepts ‘original’ and ‘variant’. These concepts are ideologically burdened since they imply that there is always one original and all other versions of the same ‘song’ are considered as variants of a known or implied original, while in fact one must actually work with concepts and ideologies that permit one to understand something as multiform (Lord 1960).

For this reason the term ‘vocal expression’ has been suggested as an alternative to ‘song’ (Lundström 2010). This term can be used globally and is not limited to one culture. Thus in the English language concepts like ‘song’, ‘declamation’, ‘recitation’, ‘incantation’, ‘chant’ are all vocal expressions. Many – perhaps most – cultures lack a meta-term for vocal expression and instead use different names for more or less distinctly different forms of vocal expressions. This is the case, for example, with the Suyá in Amazonas (Seeger 1987) and the Kammu in Laos (Lundström 2010). These different vocal expressions sometimes represent a scale from ‘speech’ to ‘song’ (here and in the following used in the general sense) and intermediate levels. They may be seen as levels of ‘speech’ or levels of ‘song’.

Among ethnic minority groups in Southeast Asia, the existence of such levels is closely related to the re-creation, extemporization or improvisation of vocal expressions. This is realized in performance by the combination of traditional sets of words or newly created sentences with pre-existing melodic and poetic templates. This is typical of the *mono-melodic* organization of music culture in which the interplay between language and music is crucial (Lundström 2010). In the special case of languages with lexical tones another interplay between language and music occurs. This can take different forms (Lundström 2010, Lundström & Svantesson 2008).

In other cultures, composition is organized in other ways. Among the Tanana Native American population in Minto, Alaska, composition is nowadays particularly done at a time before the occasion of the performance. The meter of the ‘poems’ or ‘texts’ depends on the linguistic content

and the function of the song, rather than on fixed formulas or templates (Tuttle 2011, Coray 2007). The Tanana meta-terminology applied to speaking and singing distinguishes the two in binary fashion, with subcategories of song being distinguished as types of songs, and with different categories of speaking distinguished primarily by function.

Parallels between speech and music have drawn the attention of both linguists and musicians mostly from the aspect of how and if the native language may influence music. For instance, speech rhythm and music rhythm have been shown to correlate (see Patel 2008 for an overview). In our project the same material will be approached from the different perspectives of ethnomusicology, prosody, syntax and semantics by use of their individual theoretical frameworks and working methods. Initially, we will use material from Kammu (Laos) and Tanana and Gwich'in (Alaska), languages that we have studied from these perspectives so far. The work will focus on developing a method of analysis involving and integrating the perspectives of ethnomusicology, prosody, syntax and semantics.

The Kammu and their language

Kammu (Kmhmu', Khmuʔ, etc.) is a Mon–Khmer language spoken by some 600,000 people in northern Laos and adjacent areas of Vietnam, Thailand and China. There are three major dialects, Northern, Western and Eastern Kammu. The Yùan variety of Kammu treated here is a sub-dialect of Northern Kammu, spoken in the Nalè area in the southern part of Luang Namtha Province. Northern and Western Kammu have developed a system of two tones, high and low (denoted ´ and `, respectively), while Eastern Kammu, spoken further to the east and south in Laos, and also in Vietnam, retains the original state without distinctive lexical tones. The tones have developed from voiceless and voiced initial consonants, which gave rise to high and low tone, respectively. For example, the Eastern Kammu minimal pair *klaaŋ* 'eagle' ~ *glaaŋ* 'stone' with voiceless vs. voiced initial consonant corresponds to Northern Kammu *kláaŋ* ~ *klàaŋ* and Western Kammu *kláaŋ* ~ *k*^{h}*làaŋ* with high vs. low tone. Similarly, the Eastern Kammu pair *hraaŋ* 'tooth' ~ *raaŋ* 'flower' with voiceless vs. voiced initial sonorant corresponds to Northern and Western Kammu *ráaŋ* ~ *ràaŋ;* see Svantesson (1983) and Svantesson & House (2006) for more information on Kammu tones and tonogenesis. Consonant correspondences between the Kammu dialects are summarized in Table 1.

Eastern Kammu	Northern Kammu		Western Kammu	
p t c k ʔ	p t c k ʔ	H	p t c k ʔ	H
b d j g		L	p^h t^h c^h k^h	L
p^h t^h c^h k^h	p^h t^h c^h k^h	H		H
s h	s h	H	s h	H
ʔm ʔn	ɓ ɗ	H	ɓ ɗ	H
m n ɲ ŋ	m n ɲ ŋ	L	m n ɲ ŋ	L
hm hn hɲ hŋ		H		H
l	l	L	l	L
hl		H		H
r	r	L	r	L
hr		H		H
w j	w j	L	w j	L
hw hj		H		H
ʔw ʔj	ʔw ʔj	H	ʔw ʔj	H

Table 1: Consonant correspondences between Kammu dialects (H = high tone, L = low tone).

Words and tones in Kammu

Like many other Mon-Khmer languages, Kammu has two kinds of syllables, usually called major and minor syllables. Minor syllables are unstressed pre-syllables without any phonemic vowel. Most words are either monosyllabic, consisting of one major syllable (e.g. *tís* 'mushroom') or sesquisyllabic, consisting of one minor and one major syllable, e.g. *kmùmú* [kə̀m.múʔ] 'human being, Kammu' and *kmúul* [kə́.múul] 'silver'. In careful pronunciation, minor syllables are pronounced with a schwa vowel [ə], not indicated in the phonemic transcription.

Like major syllables, minor syllables also carry a tone in the tonal dialect. In most cases, the minor syllable tone can be predicted from its segmental composition. There is a potential minor syllable tone contrast only when the minor syllable consists of two consonants and has an unaspirated stop as onset. Examples of minimal pairs are *pŋ́kà* 'to wear by the ear' vs. *pŋ̀kà* 'shy' and *pŕnɔ̀* 'broom' vs. *prǹɔ̀* 'carrying-sling'. Only about 30 minimal

pairs have been found, whereas there are more than 900 minimal pairs (such as *ráaŋ* ‘tooth’ vs. *ràaŋ* ‘flower’) for the major syllable tone. See further Svantesson (1983, 2004) and Svantesson & Karlsson (2004) for minor syllables and tones on them.

Intonation in non-tonal and tonal dialects of Kammu

In the project *Separating intonation from tone*, supported by the Swedish Research Council, we investigated how the intonation systems differ in the non-tonal and tonal dialects, based on data collected in northern Laos in November 2007 and in northern Thailand in March 2008. A total of 24 speakers ranging in age from 14 to 72 years were recorded. We found that the main function of intonation in Kammu is to signal prosodic phrasing. The words in an utterance are grouped into phrases, and the end of each phrase is marked by a high (rising–falling) pitch (boundary tone). New information is placed utterance-finally, coinciding with the place of the utterance-final boundary tone. This intonation pattern is found for both the tonal and the non-tonal dialect. The lexical tones in the tonal dialect may interfere with the realization of the intonation pattern, but do so only when the identity of a lexical tone is jeopardized. Figure 1 is a schematized illustration of the intonation pattern (smooth curve) common for all Kammu dialects (see Figures 2 and 4 for actual fundamental frequency curves of Kammu utterances). Short horizontal bars show how the utterance intonation adapts to lexical tones. The high (rising–falling) phrase-final boundary tone is changed to low only when the combination of lexical tones in the final words is high–low, since the identity of the low lexical tone would be blurred by a high phrase tone.

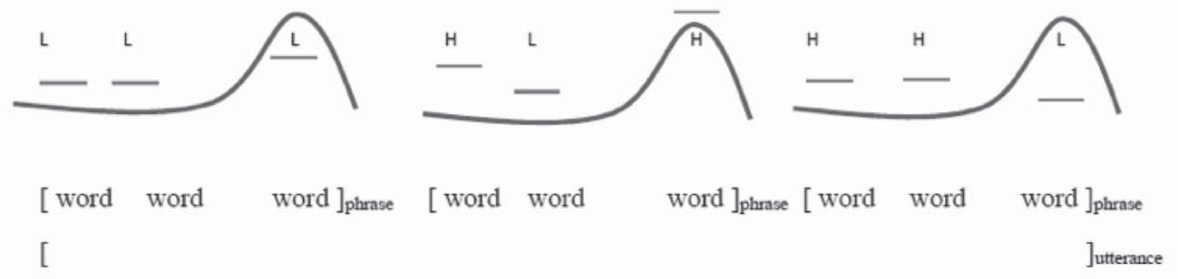


Figure 1: Stylized intonation pattern in Kammu. The curved lines show the dominating intonation pattern in both dialects. The shorter horizontal lines show how this pattern may be perturbed due to the influence of the lexical tones (marked as H(igh) and L(ow)); in particular, a phrase-final high–low lexical tone pattern (as in the final phrase) may annihilate the final rise.

To conclude, the intonation system of Kammu is quite simple with a rising-falling pitch at the end of each phrase. The presence of a lexical tone contrast in the tonal dialect forces the intonation system to adapt so that this lexical contrast can be upheld, but the adaptation is only minimal, allowing

basically the same intonation system in both dialects. See Karlsson, House & Svantesson (2012) for more details on Kammu intonation.

Alaskan Athabascan speech

There are eleven Athabascan languages spoken in Alaska; the two represented here are Tanana (Minto dialect) and Gwich'in (Venetie), both of which are considered part of the Central Alaska-Yukon subgroup of Northern Athabascan (Mithun 1999: 346). The Tanana people live on the middle stretch of the Tanana River, which flows into the Yukon. The Gwich'in people live in and near the Brooks Range and sections of the Cordillera, as well as in the valleys of the middle Yukon and lower Mackenzie (Slobodin 1981). Venetie is on the Chandalar River, a tributary to the Yukon.

For reasons of brevity, we do not include phonemic inventories from Tanana or Gwich'in, but full phonemic inventories for Tanana is given in Table 2. In Minto Tanana, laryngeal distinctions in stops and affricates are found only syllable-initially, and the voicing of fricatives and sonorants is generally predictable from environment.

		Labial	Apico-alveolar	Lateral	Inter-dental	Alveo-lar	Retro-flex	Alveo-palatal	Velar	Glottal
Stops, Affricates	Plain		t	tɬ	tθ	ts	tʂ	tʃ	k	ʔ
	Aspirated		tʰ	tɬʰ	tθʰ	tsʰ	tʂʰ	tʃʰ	kʰ	
	Glottalized		t’	tɬ’	tθ’	ts’	tʂ’	tʃ’	k’	
Fricatives	Voiceless			ɬ	θ	s	ʂ	ʃ	x	h
	Voiced			l	ð	z	ʐ	ʒ	ɣ	
Sonorants		m/b	n					j		

Table 2: Consonant inventory of Tanana Athabascan (Minto dialect)

In the Athabascan language family, tone from historical laryngeal codas varies in reflex in the modern languages: it can emerge as high or low, but in a number of languages did not develop at all (Krauss 2005 [1979]). Both the Minto dialect of Tanana, and all of Gwich'in, have low tone on some syllables, from this proto-Athabascan source. In Minto, non-low tones are unmarked for tone, and their pitches are derived from an intonational system (Tuttle 1998). The ratio of low-marked syllables to unmarked syllables – those that get their pitch from intonation – is very low, approximately 1/10 in a narrative count (Tuttle 2012). A few syllables are marked by high-rising tone. In Gwich'in, intonation has not previously been studied, and it is not clear whether non-low tones are marked or unmarked. Comparisons used for this study may provide some clues.

The intonational system of Minto Tanana is characterized by final lowering at intonational phrase boundaries, often but not always with measurable pause. Intonational low tones are lower than low lexical tones (Tuttle 1998), usually registering the lowest pitch within an intonational phrase. Utterance types (statements, questions, etc.) are not distinguished by intonational contours, but by morphological marking. There is a tendency for prefinal syllables to dissimilate in pitch from the final intonational low, and lexically low-toned syllables can be expressed with high pitch under these conditions, but this is optional. Emphasis can be expressed with high pitch, and this type of intonational effect can also override the expression of lexical tone. Intonational domains smaller than the intonational phrase have not been demonstrated for this language, but are likely to exist given findings for closely related Dena'ina (Lovick & Tuttle 2012.)

The interaction between lexical and intonational tone patterns in Minto Tanana makes lexical tone challenging to hear and to learn. Moreover, because of rich inflectional morphology and grammatical patterns, lexical tone bears almost no functional load; it is simply part of the pronunciation of words. However, we find that lexical tones are recognized in more speech-like forms of music, as will be shown below.

Vocal genres in Kammu culture

When the different vocal genres in the Kammu material are ordered on a scale spanning from 'speech' to 'song', it becomes evident that one needs to think in terms of degrees, or levels, or modes of speech – or degrees, or levels, or modes of song for that matter. Seeger (1987) arrived at the descriptions of the two extremes and an intermediary mode shown in Table 3. Seeger studied the Suyá people in Amazonia and his conclusions are here adapted – or generalized – so as to be relevant to Kammu practice.

Speech	Priority of text over melody; text and melody determined by speaker; increasing formalization in public performances.
Intermediary	Relative priority of relatively fixed texts over relatively established melodies.
Song	Priority of melody over text; time, text and melody fixed.

Table 3: The speech–song continuum based on Seeger (1987).

In Kammu tradition there are a number of rather distinct ways to perform vocal expressions. These ways – or techniques, really – have been called

vocal genres. Traditionally they were used in rather specific contexts. As is normally the case, such genres and their names may vary between areas, dialects and even individuals. The ones used in Table 4 stem from the Kammu Yùan dialect area. All except the one called *lɔ̀ɔŋ* have been studied in detail (Lundström 2010).

Lɔ̀ɔŋ	Vocal genre for reciting for example shaman songs
Hrl̃ɨ̀	Vocal genre for *trnə̀əm* with set melody template built almost totally on lexical tones.
Hrwə̀, *Húu-wə̀*	Vocal genres for *trnə̀əm* with set melody template built on variants of a genre-specific melodic motif.
Yàam	Vocal genre for *trnə̀əm* with set melody template built on a genre-specific melodic motif.
Tə́əm	Vocal genre for *trnə̀əm* with set melody template built on a melodic template that is fairly flexible and may be modified.

Table 4: The speech–song continuum in Kammu (Lundström 2010).

It must be considered, of course, that the listing of these vocal genres on a scale from speech to song should be seen as a suggestion. Despite some overlapping and other inconsistencies it provides, however, an overview that is in many ways useful. It is also possible to see a continuity within the poetic techniques that are used in the different vocal expressions and to list them from simple to complex. In this chain each poetic technique may be seen as a special and more complex case of the preceding one:

Reduplication → Lexical parallelism → Pivot-rhyme → Chain-rhyme → Cross-rhyme

Table 5 shows vocal genres in a continuum from ‘speech’ to ‘song’ combined with the poetic techniques as they tend to appear. The vocal genres that will be discussed in the present study are narrating, *lɔ̀ɔŋ* (in the case of a Guiding song), *hrlɨ̀* and *tə́əm*.

Speech Song

Vocal genres:											
Daily speech	Formal speech	Narrat-ing	Indirect speech	Prayers	*Lɔ̀ɔŋ* (recite)	Children's songs	Ceremonial songs	*Hrlɨ̀ɨ*	*Hrwə̀, ˎ Húu-wə*	*Yàam*	*Táəm*
Poetic techniques:											
Reduplication											
Vowel or consonant rhymes (Sayings, Proverbs)											
	Pivot rhymes Lexical parallelism										
			Chain-rhyme								
								Cross-rhymes (*trnə̀əm*)			
											Words of ad-dress

Table 5: The speech–song continuum in relation to poetic techniques.

Vocal genres in Alaskan Athabascan culture

The two groups discussed in this paper, Tanana (Minto dialect) and Gwich'in (Venetie dialect), have developed differently with respect to music and dance since contact with white culture. The Gwich'in encountered Europeans (French and English speakers) in the 19th century and developed their own Western musical styles, adding violin and guitars to their traditional vocal and drum tradition (Honigmann 1981: 732). Their older traditional music is less practiced today than that of the Tanana, who have held on to a strong native musical tradition up to the present generation. The present generation of elders in Minto is now a very small group of fluent speakers and singers, and younger people are not acquiring the language, although they are studying the musical tradition. Gwich'in consulted for this study say that they have not heard much traditional singing, but there are more young speakers of Gwich'in than of Tanana. Accordingly, data for Tanana includes recordings made in the last five years, while the Gwich'in data was recorded in 1972, and archived at the University of Alaska Library as part of a 'Songs and Legends' collection.

The traditional music of the Tanana has been studied since the 1960s by musicologists (Lundström 1980, Pearce 1985, Johnston 1994), who divide the songs into various categories, based on function (Lundström) or source (Pearce) or some combination of these two variables (Johnston). Since all these divisions accompany differences in style, the approach here is to sort the data according to structure more than function or source. Categories of vocal expression have been defined based on interviews with native

speakers and practitioners of vocal art, and these categories are reflected in terminology within the language.

We consider a dance song, two mourning songs, a medicine song, and one personal or art song with mixed Athabascan and English lyrics, a post-contact native style found in both groups. Some of these songs have been previously discussed in Tuttle (2011). Thanks are due to Neal and Geraldine Charlie, Sarah and Bergman Silas, Susie Charlie, Kathy Sikorski and Allan Hayton for their help with translations and explanation.

In the analysis these songs will be studied with the combination of language and music as the fundamental approach and with the ultimate aim to investigate whether in Athabascan culture there is something similar or parallel to the templates of the vocal expressions found in Kammu culture. This will be a first step in this direction.

Prosodic comparison of Kammu vocal genres

At the present stage we will put forward some aspects of rhythmic and tonal patterns of spoken Kammu that may be relevant for investigating transitions within the speech–song continuum, namely segmental and tonal rhythm.

Both speech and song use variations in pitch and length (duration in time) of certain units, and the reasons for and functions of these variations are different for speech and song. In speech, length can be used for distinguishing lexical meanings, for example by long and short vowels in Kammu. In other languages, alternations in length can be used to mark strong syllables by longer duration (e.g. *'increase* vs. *in'crease* in English). When pitch is used for lexical purposes (as in Kammu) we speak about tonal languages. In addition to its lexical function, pitch is also used on the utterance level, typically to mark important information by highlighting words or groups of words with e.g. higher pitch (focal accent), and to mark the end of a phrase by pitch change (phrasing). In many European languages, speakers have rising pitch at the end of questions (boundary tone) or of uncompleted utterances, and falling pitch marks the end of a declarative. Each language has a specific inventory of possible pitch combinations (intonation grammar) and they can only be used in certain context. Thus, speakers are not free to modulate their pitch at will.

It is generally acknowledged that pitch and length are also used to create speech rhythm, and variations in pitch and length are perceived as variations in speech rhythm. Speakers of the same language can shift between higher or lower degrees of rhythm by using tonal means, as is shown for Kammu in the next section. Thus, pitch and length can be used phonologically to distinguish meaning and non-phonologically for highlighting, phrasing and rhythmization. One might assume that phonological (contrastive) features tend to be preserved in the transition from speech to song, influencing the

melodic structure of songs. Non-contrastive pitch and duration are used to create speech rhythm, however, and may because of this influence song rhythm as well.

Like many other languages (such as Finnish, Mongolian and Swedish), Kammu uses vowel length for lexical contrasts. Thus, we find minimal word pairs as *már* 'barren' vs. *máar* 'salt', differing only in vowel length. Long and short vowels contrast in major syllables that end in a sonorant (*m, n, ɲ, ŋ, l, r, w, j*) or a stop (*p, t, c, k*), but there is no vowel length contrast in syllables ending in *h, ʔ* or in open syllables. Are these length contrasts preserved in song? Do prolongation patterns depend on whether a language uses length phonologically or not, or whether it has lexical stress (as English and German) or lacks it (as Japanese and Korean)? Asano & Braun (2011), for instance, found different patterns of prolongation in German and Japanese speech. Are there similar language dependencies in song? Previously, we found that words (major syllables) are prolonged in Kammu speech when a word is emphasized, but the phonemic contrast between long and short vowels is still preserved (Karlsson, House & Svantesson 2012). In *hrlɨɨ* we find patterns different from speech but still connected to phonological contrasts (see below).

Regarding prosody, transitions along the speech–song continuum can be assumed to be made by two main principles: changes in segmental and tonal rhythm (lower or higher rhythmization) and creation of aesthetic contrasts, e.g. by using prosodic patterns that only occur in one kind of vocal genre. We will show that the speech–song continuum cannot be established on prosodic grounds for Kammu. Instead, transitions between speech and other genres are built by choosing some particular prosodic features creating independent transitions: speech > narratives > Guiding songs; speech > tale telling; speech > *hrlɨɨ,* etc.

Prosodic transitions between vocal genres in Kammu

The speech material used for this initial description are recordings of read speech of tonal Kammu (sentences of different length translated by Kammu speakers from Lao or Thai into Kammu) and recordings of the same speakers telling how rice is grown and prepared, from the making of a rice field until the rice is eaten (henceforth 'rice narratives'). 14 tonal speakers were recorded in Laos in 2007 and in Thailand in 2008. Recordings of two tonal speakers telling a folktale, recorded by Kristina Lindell in the 1970s, are also used here. The song material chosen for this introductory study comprises one performance of *hrlɨɨ* by a tonal speaker, one performance of the Guiding song (*Lɔ̀ɔŋ ŋɔ̀ɔr*) spoken by one tonal speaker, and one performance of the Guiding song sung by tonal speakers (leader and choir).

When comparing speech genres, a high degree of rhythmization is found in the rice narratives. Rhythmization is attained by tonal, segmental and syntactic means. Very few disfluencies (interruptions, hesitations, etc.) are found, as speakers are telling about growing rice, an everyday activity well-known for most of them. Although spontaneous, this material is structured in the same way by all speakers: new information is given at the end of an utterance (underlined in (1)) and is then repeated in the immediately following utterance (see further Karlsson, House & Svantesson 2012):

(1) [hóoc nì yɔ̀h kíaw] [kíaw ŋɔ́] [kíaw hóoc yɔ̀h tíi]
[finish then go harvest] [harvest rice] [harvest finish go beat]

[tíi ɔ̀ɔr wèc] [tàa yùuŋ]
[beat bring return] [to barn]

Figure 2 shows the tonal course of this sentence spoken by a female tonal speaker. Major phrasing is shown by vertical lines. Tonal boundary tones of focused phrases, occurring on the underlined words in (1), are shown with arrows.

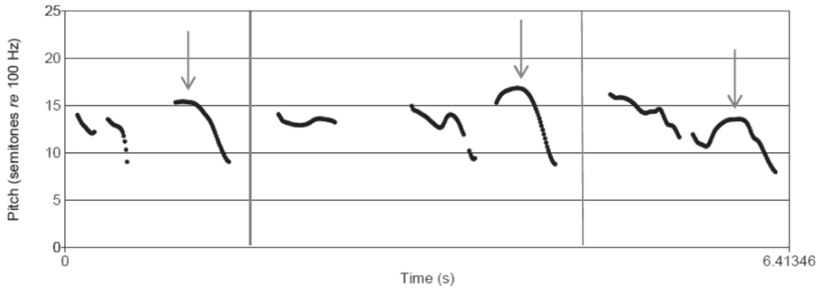


Figure 2: Tonal course of utterance (1).
Sound file PhA-JB2013_KLST_Fig2_rise_narrative.mp3

Anchoring with the preceding content is achieved by repeating parts that have been mentioned immediately before, for example: 'When the rice is ripe, we harvest it. After we have harvested the rice, we carry it home to the village'. Such lexical repetitions are reflected in higher tonal and segmental rhythmicity as shown in Figure 3. In this illustration, word length is depicted by boxes, and falling boundary tones by schematic curves. It can be seen that the boundary tone is much higher and the word length is greater in focused positions than in unfocused ones.

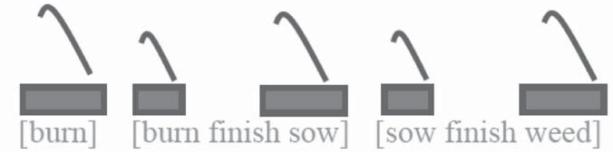


Figure 3: Schematic illustration of segmental and tonal rhythmicity in the rice narratives.

Interestingly, a similar syntactic structure, with repetitions, is found in the Kammu Guiding song *Lɔ̀ɔŋ ŋɔ̀ɔr,* which was sung or recited at funeral wakes

in order to guide the soul of the deceased to the spirit village where s/he will live as an ancestor spirit (Tayanin 2006). In our material, the Guiding song is performed in two ways, as the vocal expressions *lɔ̀ɔŋ* or *yàam* (cf. Table 4), and when performed by more than one person it is performed in a call/response manner. The analysis here is based on the *lɔ̀ɔŋ* versions. This is an example of a portion of the text of the Guiding song:

(2)

Cə̀h	Klàaŋ	Ktéey	cùur	ə̀nsú	rɔ̀ɔt	Òm	Críil.
from	stone	*Kteey*	go down	down	arrive	water	golden.

'Go down from the Ktéey cliff to the Golden Brook.'

Cə̀h	Òm	Críil	lɔ̀ɔŋ	ə̀nsú,	rɔ̀ɔt	Wàk	Kntɔ̀.
from	water	golden	follow	down,	arrive	waterfall	rainshield.

'Go along the Golden Brook down to the Rainshield Waterfall.'

Kháam	Wàk	Kntɔ̀	ə̀nsú	rɔ̀ɔt	Òm	Cùk.
pass	waterfall	rainshield	down	arrive	water	fishtail palm.

'Go past the Rainshield Waterfall down to the Fishtail Palm Brook.'

Figure 2 shows the tonal course of this sentence spoken by a female tonal speaker. Major phrasing is shown by vertical lines. Tonal boundary tones of focused phrases, occurring on the underlined words in (1), are shown with arrows.

The Guiding songs can be spoken or sung. Thus, we can study similarities and differences between highly rhythmic everyday narratives (rice monologues) and spoken or sung Guiding songs. The initial part of the Guiding song is structurally very similar to the rice narratives: new information is given in the second phrase and this phrase gets a higher tonal boundary. As the speaker gets into the rhythm, the tonal structuring is changed (see Figure 4, top panel). The main principle is still the same: the last word in a phrase gets a high boundary tone. Here, however, we find that the phrases with old (repeated) information (the first and the third ones) are realized at a higher pitch level than the phrases comprising new information (the second and the last phrases). This pattern is typologically somewhat unusual in speech, as prosodic marking of new information is usually associated with a higher articulatory effort and because of this with a higher pitch. A similar pattern is found in sung Guiding songs, as illustrated in the bottom panel of Figure 4: the right phrase boundary is marked by a high boundary tone (suppressed by a low lexical tone in the second phrase), and phrases comprising new information (the second and the last one) are realized with lower pitch.

Thus, we find a higher rhythmicity in speech in narratives and their prosodic and syntactic features are very similar to Guiding songs. Tonal

phrasing and the use of lengthening are similar in the two genres, and the main difference is the use of pitch levels. In the Guiding songs, the speaker shifts between higher and lower pitch levels (henceforth registers) by performing every second phrase on a lower level. In speech, the speaker realizes all phrases within the same register, shifting the magnitude of the peaks of boundary tones, so that an upstepping of the peaks is observed. Realization of the whole phrases within three different registers (extra low, low and high) is found in other song genres for tonal speakers (Lundström 2010), and preliminary observations of songs of non-tonal speakers do not indicate such differences in registers. We will study these differences in more detail as they may tell us more about how lexical tones are preserved in songs and how non-tonal Kammu speakers differ in song production. The possible use of pitch registers in the speech of tonal speakers will also be investigated in more detail.

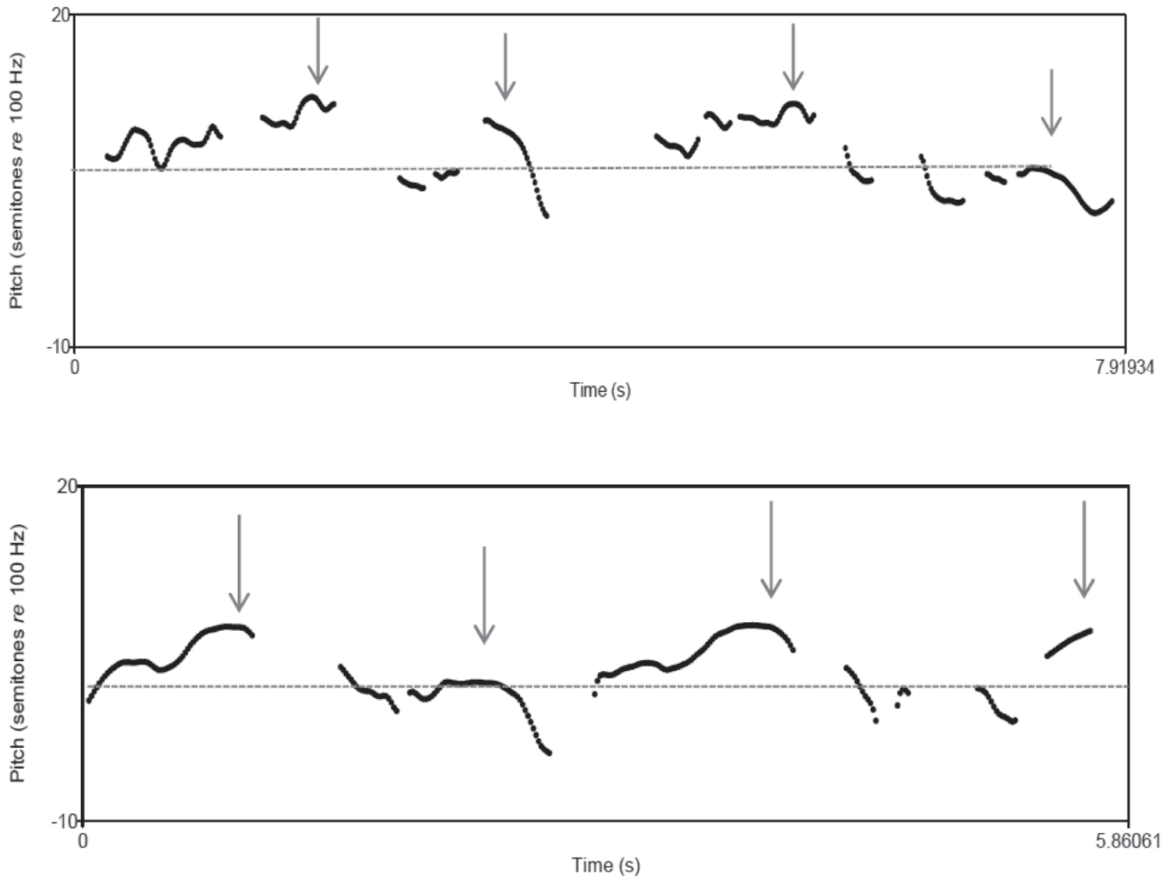


Figure 4: Tonal course of a later part of a spoken Guiding song (top) and of a fragment of a sung Guiding song (bottom); male tonal speakers. Ends of phrases are shown by arrows and the horizontal line indicates an approximate division into high and low pitch level. Sound files PhA-JB2013_KLST_Fig4a_guiding_spoken.mp3 and PhA-JB2013_KLST_Fig4b_guiding_sung.mp3 respectively.

Compared to narratives and Guiding songs, storytelling has another type of rhythm and some parts are perceived as less rhythmic than narratives. It has, however, a very interesting pattern not found in everyday speech. While lexical tones and length contrasts seem to be preserved, speakers use phrasing, emphasis, voice quality and prolongation as aesthetic prosodic means in storytelling. The phrasing pattern is not as reoccurring as in everyday speech and narratives, and emphasis of some non-final words or groups of words occur, unlike in everyday speech. Some words are considerably prolonged, but length contrasts are still preserved. Function words (such as demonstratives and interjections) are free regarding lexical tone and length. These are however very first observations.

The Kammu genre *hrlɨɨ* in the tonal dialect

In the vocal genre *hrlɨɨ* in tonal Kammu, the lexical tones are clearly dominating. The ratio between musical pitch dominance and lexical tone dominance can be approximated to 10:90 (see further Lundström & Svantesson 2008). *Hrlɨɨ* singing is strictly syllabic, i.e. each tone in a song corresponds to one syllable. For the main part it employs only one tone duration – the only exceptions are the penultimate syllable of a line and the very last syllable of a stanza, which are longer. The relative duration pattern for the two first lines is: [short – long – long – prolonged – short, short – long – short – long – long – prolonged – long]. With these exceptions, each syllable is given the same length regardless of vowel length. A minor syllable is, without exception, treated in the same way as a major syllable and is thus given a much longer relative duration than in speech. An example of a *hrlɨɨ* song is given in (3), and in Figure 5, the same song has been reorganized so that the minor syllables and initial formulae are cut out and the lines of the poem have been re-grouped to show consistencies in the intonation of lexical tones (H = high, L = low lexical tone).

(3) àay màh krɛ̀ nɔ́ɔŋ ɔ̀ɔn
krɛ̀ nɔ́ɔŋ ɔ̀ɔn pùun-pùun kl̀-táak
àay màh krɛ̀ nɔ́ɔŋ ɔ̀ɔn
krɛ̀ nɔ́ɔŋ ɔ̀ɔn pùun-pùun tŋ̀-kɨ́l

àay màh kɔ́ɔn nɔ́ɔŋ nɛ̀
kɔ́ɔn nɔ́ɔŋ nɛ̀ hń-tàak ḱ-núun
àay màh kɔ́ɔn nɔ́ɔŋ nɛ̀
kɔ́ɔn nɔ́ɔŋ nɛ̀ rŋ̀-kɨ̀l ḱ-núun

Translation:

I am still weak like a plaited table,
Weak like a plaited table, like a stepped-on tree-trunk.
I am still weak like a plaited table,
Weak like a plaited table, like a stepped-on tree-stump.

I am still small like a little child,
Small like a little child, less than knee-high.
I am still small like a little child,
Small like a little child, just about knee-high.

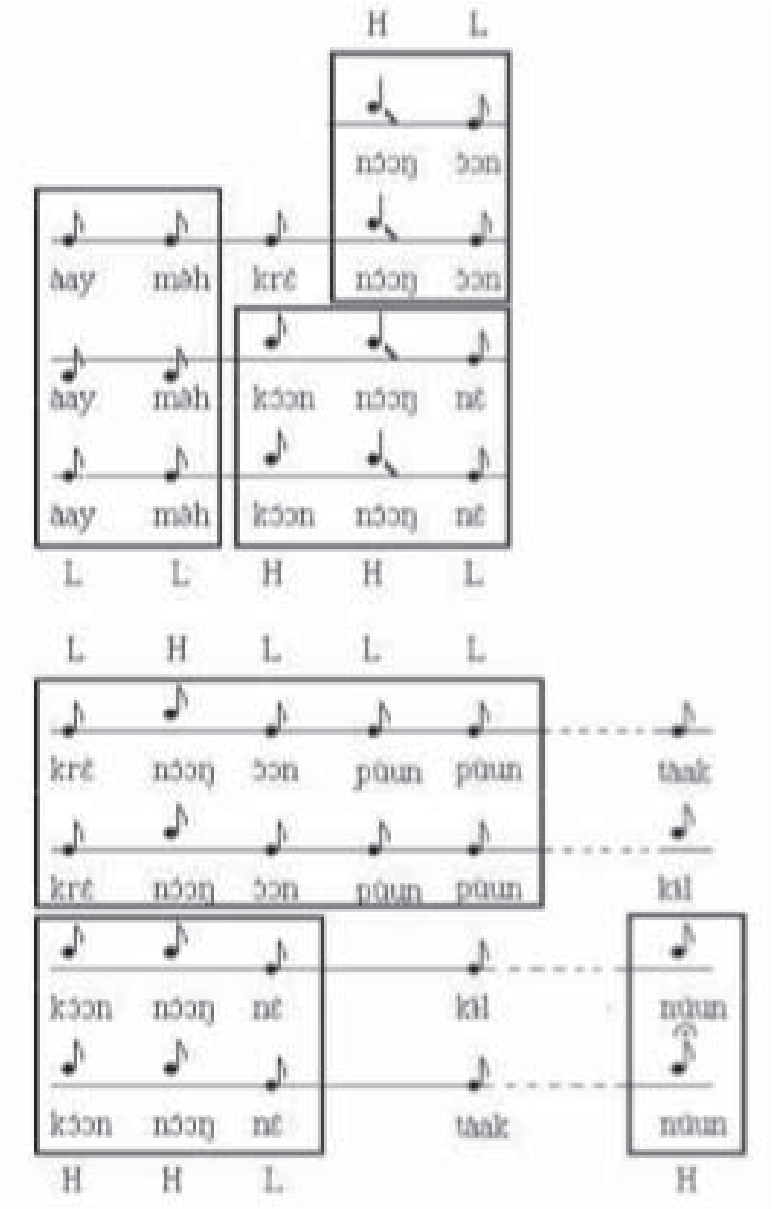


Figure 5: Reorganized transcription of the ***hrlɯ̀*** song (3)
where approximate pitches are indicated.
Sound file PhA-JB2013_KLST_Fig5_hrlii-aay-meh.mp3

In this particular example, it is not possible to tell whether length contrasts are preserved since words with non-long vowels in this song are of non-contrasting type. What we can observe is that prolongation takes place in syllable-final sonorants (*nɔ́ɔŋ* 'still' > *nɔɔŋŋ*) and by inserting epenthetic vowels in minor syllables and prolonging them (*rŋ.kɨ̀l* 'no more than' > *rəəŋŋ.kɨ̀l*). Thus, prolongation is achieved using positions in which the length contrast does not occur. This contrasts with the situation in normal speech where the minor syllables are unstressed and have a very short non-phonemic epenthetic vowel which often disappears completely. In *təə́m* singing, the vowel of the major syllable can reduplicate and build a new (long) syllable, e.g. *mɛ̀ɛn* > *mɛɛ.nɛ* or *mɛ.nɛɛ* (Lundström 2010).

For the main part of the performance, *hrlɨ̀ɨ* singing is limited to two pitch levels – the only exception being a few syllables at the beginning of a poetic line after a pause, which are sung to an extra-low tone and can be considered as an introductory formula. The size of the interval between the two dominating pitches varies from a 2nd to a minor 3rd, which means that the pitches can be easily recognized by ear. As will be seen below, the high and low singing pitches are almost invariably used for high and low lexical tone respectively. 'H' and 'L' stand for high and low lexical tone respectively, in Figure 5, while the relative musical pitches are shown by the notes. Repetitions of words in this performance are grouped in boxes to demonstrate that the same lexical tones are performed in the same way each time they occur.

In the investigation reported in Lundström & Svantesson (2008), we used material consisting of performances by one Kammu speaker, Kàm Ràw (Damrong Tayanin), belonging to the Yùan subdialect of the Northern dialect. A 'studio sample', which was sung on the informant's own initiative, consisted of 12 performances and 909 syllables in total. A 'laboratory sample', which was sung on the request of the researcher (HL), consisted of 24 performances and 1,393 syllables. Finally, there was an 'experiment sample' consisting of another 12 performances that the informant had not sung in this style before. This was made in order to test predictions about rhythm and pitch relative to lexical tones. A fourth sample consisted of 2 performances by other Yùan informants. In total, then, there were 50 performances. A musical transcription of all songs was made by one of the authors (HL).

The results showed that the *hrlɨ̀ɨ* material followed the linguistic analysis of both syllabification and of the individual tones, both on major and minor syllables. There were no mismatches at all in syllabification and the tonal mismatches were extremely few, less than 0.5%.

The Kammu genre *tə́əm* in a tonal dialect

Tə́əm singing is an elaborate form of performing the orally transmitted poems *trnə̀əm*. The singer may elaborate also the words by prolonging or contracting them or by adding sets of words that may be traditional or made up on the spot. A *tə́əm* melody basically consists of a melodic template that varies between the Kammu dialect areas, but also in details between villages and even between individuals. Such a template has rather determined initial and final formulae sung on various words that may be translated 'hey' or 'oh'. A common final formula is *kàay sáh*, 'this I say', and the word *sáh*, '[I] say', is often squeezed in at the start of lines within the *trnə̀əm*. Apart from these portions, the melodic template moves forward in a sort of recitation where the lexical tones are more or less slavishly adhered to. A main question for future research is the relation between melody and words in *tə́əm* singing in the non-tonal Kammu dialect.

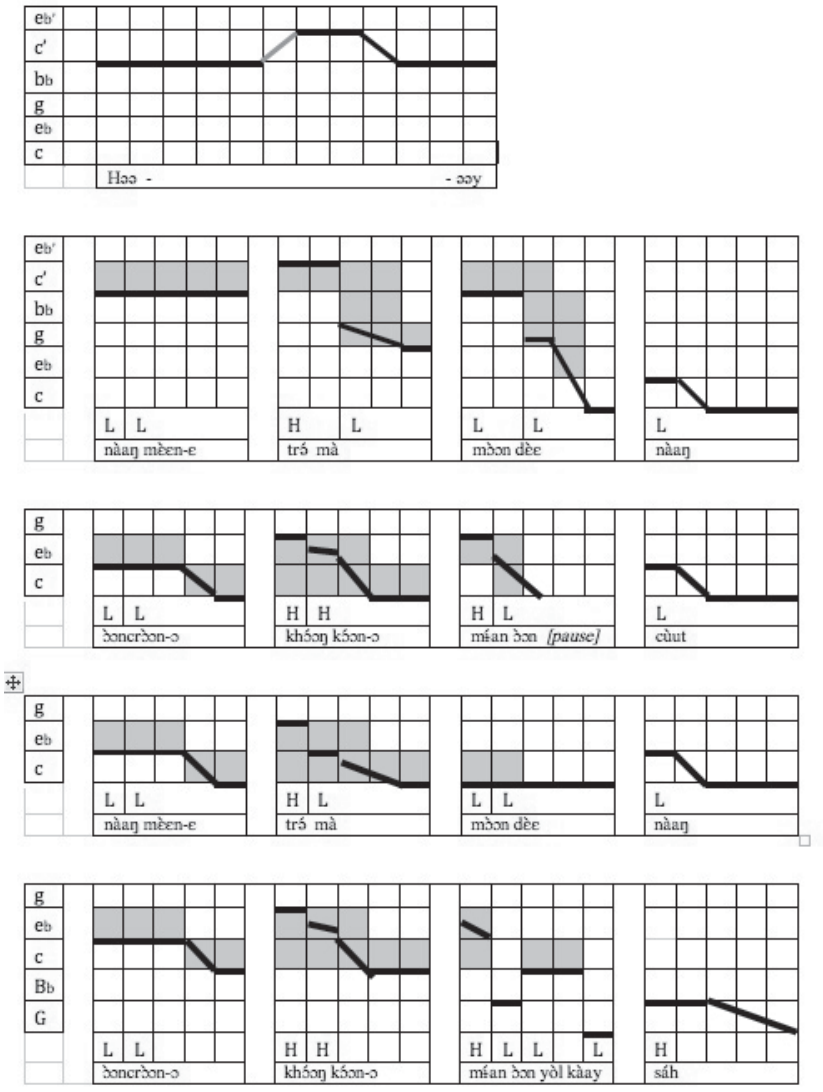


Figure 6: One phrase of a *tə́əm Yùan* performance.
Sound file PhA-JB2013_ KLST_Fig6_teem-naang-meen.mp3

For those parts of performances where lexical tones and musical pitch coincide we use the term *lexical tone centration* whereas the other parts are called *music pitch centration*.[1] When initial and final formulae are subtracted, i.e. the parts where music pitch centration dominates, a performance is more influenced by lexical tone centration. For large parts of the recitation the words are grouped two by two and performed in an iambic pattern (short + long). The correspondence between lexical tone and musical pitch is at least 50 per cent. However, this also means that lexical tone recitation is not strictly applied. For example, the final syllables are always sung in the same manner, regardless of the lexical tone. Figure 6 is an approximate graph of one phrase of a *tə́əm Yùan* performance. The shaded parts show the areas where words can be predicted to be sung if the lexical tones were different from this particular performance.

A discussion of Alaskan Athabascan songs

The study of vocal genres in Alaskan Athabascan languages is still developing, building on investigations of intonation. One aim with our study is to attempt to understand composition or re-creation and how words and music are interrelated in these processes. This will in turn hopefully make it possible to understand Alaskan Athabascan vocal expressions from their structural characteristics.

To date, intonational findings for Alaskan Athabascan languages have demonstrated patterns that are common to other languages, including English: final lowering (Tanana, Tuttle 1998; Dena'ina, Lovick & Tuttle 2012) and lengthening at the ends of intonational units, reset at the beginnings of intonational units (Ahtna, Berez 2011), and manipulation of timing and pitch in narrative (Lovick & Tuttle 2012). Tuttle (1998, 2003) finds that the lexical low tone in Tanana is acoustically distinct from final lowering. Holton (2005) shows that final lowering has important effects in Tanacross as well; since this language has high tones where Tanana has low tones, the overall effect is different, but the phenomenon of final lowering is the same.

The relationship between intonation and tone varies among the Alaskan Athabascan languages. For obvious reasons, languages with no tone (Dena'ina, Ahtna) have been chosen for base studies of intonation in narrative (Lovick & Tuttle 2012). Studies of languages with tone (Tuttle 1998, 2003, Holton 2005) show somewhat different relations between the two prosodic systems, but in both Minto and Tanacross, intonational patterns in speech can override the expression of the sparse tones due to

1 These terms were introduced in Lundström (2010: 48).

historical laryngealization. This means that a general contour, including an intonational-phrase-final low, can override the expression of either a low or a high tone. This does not mean that the tones are lost for fluent speakers; they are recoverable when intonational structure does not conflict (Tuttle 1998.) Thus, Tanana lexical tones are variably expressed depending somewhat on environment. There is also some variability among speakers, which roughly corresponds to their age, but can also be affected by other factors, such as contact with other Athabascan languages. The effect of historical low tone is subtle and difficult for outsiders to perceive and to learn.

Questions about vocal genre have been based on the ideas that intonation can be manipulated by speakers to identify genre, and that genres can be identified by prosodic characteristics. The former assumption appears to be borne out by research to date, but the latter hypothesis is not as easy to support, in part due to the small number of narrators represented in the audio and transcribed record. Non-musical vocal genres in the Athabascan languages seem to be defined more clearly by the structure of their content. From our research to date, we find no reason to identify Alaskan Athabascan narrative with poetry, as might be suggested by a strict application of methodology demonstrated in Hymes (1981), for example. However, we have followed established tradition in the analysis of Native American literature, in transcribing stories by prosodic grouping rather than by syntactic boundaries. We find that prosodic groupings have relatively regular patterns, which can be manipulated by speakers to create narrative effects. Narrative genres also have recognizable markers for beginnings and endings, which are lexical in nature but may be accompanied by prosodic cues. The bases for these claims come from research on non-tonal Dena'ina (Lovick & Tuttle 2012) and tonal Minto Tanana (Tuttle 2011), both of which were based on original fieldwork and the study of archived recordings.

Song lyrics are not stories, of course, and they are also not poetry in the sense that Westerners understand it – they are not even orally transmitted poetry, but structured speech inextricably linked to song. Much of the "poetry" analyzed by Hymes in his examination of published Native American materials is actually lyrical content abstracted from songs and chants. He himself states that "[t]he absence of the musical portion of the songs may admittedly be a source of error […] The printed texts, however, seem intended as lexical abstracts, not as verse structures […] In any case, the choice with the present texts is either to analyze them as texts or do nothing, since no music exists." (Hymes 1981: 43–44).

In the present case, we are much more fortunate in that we have both lyrics and music to work with, and speakers of the language who

understand the language in the lyrics and who can write songs themselves. The determination of line endings is not problematic, as the musical lines, including meaningless vocable portions,[2] demonstrate line structure through musical and linguistic parallelism (see Powers 1992 for discussion of the structural role of vocables). The music, as in the Hopi songs analyzed by David Shaul, "is an envelope, a larger structure and context into which the linguistic text fits" (Shaul 1992: 232).

The structure of Tanana song lyrics ranges from relatively prose-like to strictly metered, depending on the function of the song. Memorial songs and some medicine songs display prose-like qualities, with melody following tone and intonation to a greater or lesser degree. Dance songs are highly rhythmic, and speech tone and intonation are not expressed directly when they are sung.

We have added two songs that contain lyrics in another Alaskan Athabascan language, Gwich'in, with some code-mixing from English and a third Athabascan language, Koyukon. The prosodic structures of Gwich'in and Koyukon have been studied in far less detail than those of Tanana, but Gwich'in is classed as low-toned according to Athabascan convention. Koyukon is considered not to have lexical tone, but tonal effects may exist in some under-studied dialects.

1. *Dance Song: Joni ło'o* 'Here it is!' (Tanana).
Sound file *1-joni-lho'o.wav.*

This is a *ch'edzes ch'elik*, or 'dance song'. Dance songs are performed at gatherings of all types, including formal potlatches in honor of deceased or living persons, and may also be performed for audiences. The Tanana expression literally means 'dance song', and is a compound of two nouns and not a description. It is the term used by elders when designating this type of music. The lyrics to this song are limited to one sentence (4); see also Figure 7.

(4) *Joni* *łoʔo* *Ganhok* *tolał*
here FOC Ganhok 3sS.**be**.FUT
'Here, this will be the Ganhok.'

The *ganhok* is a tall dance staff. The song is said to have been composed on the occasion of the introduction of the *ganhok* to the Tanana people (the word is of uncertain origin, though there is speculation that the staff was introduced through Tlingit contact). While the recording here was made by

2 In this article we use the linguistic term 'vocable', which is commonly employed in leading studies on Native American music, rather than 'vocalise', which in turn is common in musicology but does not have a clearly defined linguistic meaning. Vocable, in linguistics, denotes sounds without meaning that may be used in alternation with words that do have meaning.

a small group of elders, normally this song would be performed by a hall full of people all shouting together.[3]

In this song the melody and rhythm overwhelm the words, and all the meaningful syllables are given the same weight, one strong beat. Phrase-final vocable syllables are lengthened over several beats. There are no tonal syllables in this lyric. There are eight meaningful syllables in the song, arranged in two groups of four. The following vocable sequences also pattern in groups of four. All the syllables are heavy according to the Tanana stress system, in which full vowels and closed syllables count as heavy. A sequence of four heavy syllables would not be impossible in speech, but is usually broken up by light syllables (CV where V is a schwa) in the normal composition of sentences.

The first half of the text (A, *Joni łoʼo)* is performed as three short tones and one long: _ _ _ __ and the same rhythm is repeated for the second half (*Ganhok tolał*). The final tone of these words functions as a tonic or tonal center of the song, so they finalize the first section of the song. The second section (B) is built on the same rhythm, which is repeated twice, but sung with vocables with no lexical meaning (*Iya ha e*). For the third repetition these syllables are sung at the same pitch, which further stresses the tonic.

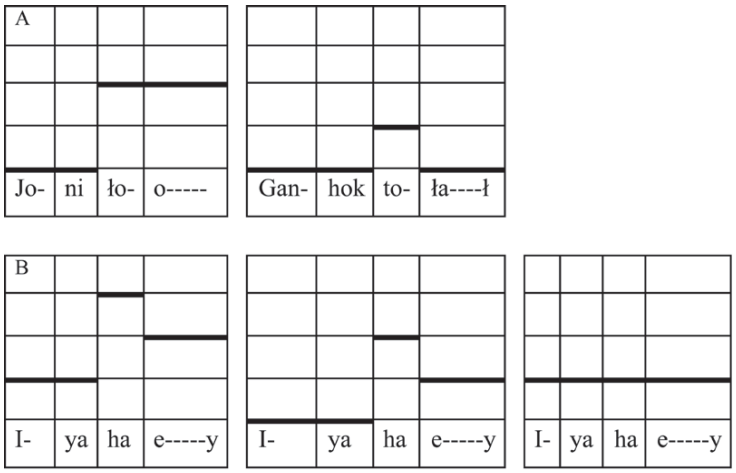


Figure 7: Dance song: *Joni* łoʼo. Sound file PhA-JB2013_KLST_Fig7_joni-lhoʼo.mp3

This song shows several characteristics that are common in Native American music: the song is strophic (the diagram in Figure 7 shows one stanza that is repeated with minor changes), it is isorhythmic (that is: built on one rhythmic idea throughout), it has a stepwise downwards or falling motion (the second section of the stanza) and it ends on a tone repetition on the lowest pitch (that serves as a tonic, denoted with a wave-line in the graph).

3 The YouTube video entitled "FNA 2012 Troth Yeddah' dancers" below gives a general idea of the presentation style, and shows dancers using the *ganhok*, though they are singing a different song: <http://www.youtube.com/watch?v=2VKcFTdl8P8>.

In his study on Tanana Athabascan dance songs, Pearce (1985) concludes that the overall form of such a song is usually binary, i.e. there are two distinct sections within each stanza of a song. He also finds that dance songs are often built around one melodic/rhythmic motif that is repeated at consecutively lower pitch, which with a musicological term is called a *sequence*. He calls this kind of motif a *building block*, which is a very suitable name (denoted by a double bottom line in the graph).

This song may represent a category of song that takes its start in a poignant melodic/rhythmic setting of a verbal phrase, which becomes the nucleus and serves as a building block, combined with vocables, sequences in downward motion and an ending with a tone repetition on the tonic.

2. Mourning Song: *Dolo K'adi* 'Missing Dolo'.
Sound file *2-dolo-k'adi.wav.*

Dolo K'adi is a song of the type called *dratax ch'elik* in Tanana and referred to in English as 'sorry songs'. *Dratax* is a noun derived from a verb meaning 'to dance a mourning dance, arms moving up and down' The term *dratax ch'elik* is a noun compound. This song was composed in the 1930s, and the composer used rather complicated 'high' language that is not easy for present-day speakers to translate. *Dratax ch'elik* are composed to honor an individual who has died, and are performed at memorial potlatches, held a year or more after the honoree's passing. They are composed by family members of the deceased, or by community composers writing a song for them to sing. The family member decides on the verbal content to be included, in the latter case, and the composer assists them in organizing the lyrics and then sets them to a melody. The melody is not always new, but must not too closely resemble important songs that may be in the repertory of the local community or a neighboring community. New compositions are evaluated by song leaders and may be revised if they notice errors. Older *dratax ch'elik*, especially really beautiful ones, are often performed at funeral potlatches (held just after a death) when there has not been time to create new songs. The song analyzed here is performed widely at women's funeral potlatches, and is known to speakers of other languages in Alaska due to kinship ties that bring Minto singers to potlatches statewide.

Dolo is the name of a woman. In most *dratax ch'elik* examples we are aware of, only one kinship term is used in the A section (e.g. *en'a* 'mother'). In the case of *Dolo K'adi,* two terms are used, 'mother' and 'sister', which emphasize the great importance of the honoree to her family. Dolo was the eldest daughter of a famous Minto chief in the early 20th century, and the singers consulted in this study are her descendants.

(5) Mourning Song: *Dolo K'adi*

A

Oh-oh-oh-oh	*En'a'a*	*oh-oh-oh-oh-oh*	*En'a'a*
VOC	**mother**-VOCATIVE	VOC	**mother**-VOCATIVE
oh-oh-oh-oh-oh	*En'a'a*	*Soda*	
VOC	**mother**-VOCATIVE	1s-POSS+**older-sister**	

Line 1 'Mother, mother, my older sister'

B

ekhwdon'a	*ch'ukat*	*dinot*
next-upriver	INDO+3+**buy**	**while**

Line 2 'Just upriver, while out shopping'

logha	*dit'a*	*khełdi*
handy	2sS+**be**+STAT	3hS+**with**+**speak**-IMPERF

Line 3 'You are handy, they say'

nelo'	*dodelu'*
2s-POSS-**hand**	**speaking**-3sS-**be-charming**-IMPERF

Line 4 'Your hands are worthy of praise.'

Ye'ał	*khenino*	*doch'edenaghiloyh*	*yino*
3+**with**	3hS+**come**-PERF	**speaking**-INDO+2sS+**gather**-IMPERF	**while**

Line 5 'And they came, while you gather things together'

En'a	*Soda*
mother	1sposs+**older-sister**

Line 6 'Mother, my older sister'

A

Oh-oh-oh-oh En'a'a, oh-oh-oh-oh-oh En'a'a, oh-oh-oh-oh-oh En'a'a Soda

C

Yełni	*khw*	*khełdi*
3sO+**with**+**speak**-IMPERF	**as**	3hS+**with**+**speak**-IMPERF

Line 7, Line 10 'As she told him/her, they say'

ekhwdon'a	*ch'ukat*	*dinot*
next-upriver	INDO+3+**buy**	**while**

Line 8 (2) 'Just upriver, while out shopping'

logha	*dit'a*	*khełdi*
handy	2sS+**be**+STAT	3hS+**with**+**speak**-IMPERF
Line 9 (3)	'You are handy, they say'	

Ye'ał	*khenino*	*lo*	doyedenaghiloyh	*yino*
3+**with**	3hS+**come**-PERF	foc	**speaking**-INDO+2sS+**preserve**-IMPERF	**while**
Line 11	'And they came to it, while you put it up [food]'			

En'a	*Soda*
mother	1sposs+**older-sister**
Line 12 (6)	'Mother, my older sister'

A
Oh-oh-oh-oh En'a'a, oh-oh-oh-oh-oh En'a'a, oh-oh-oh-oh-oh En'a'a Soda

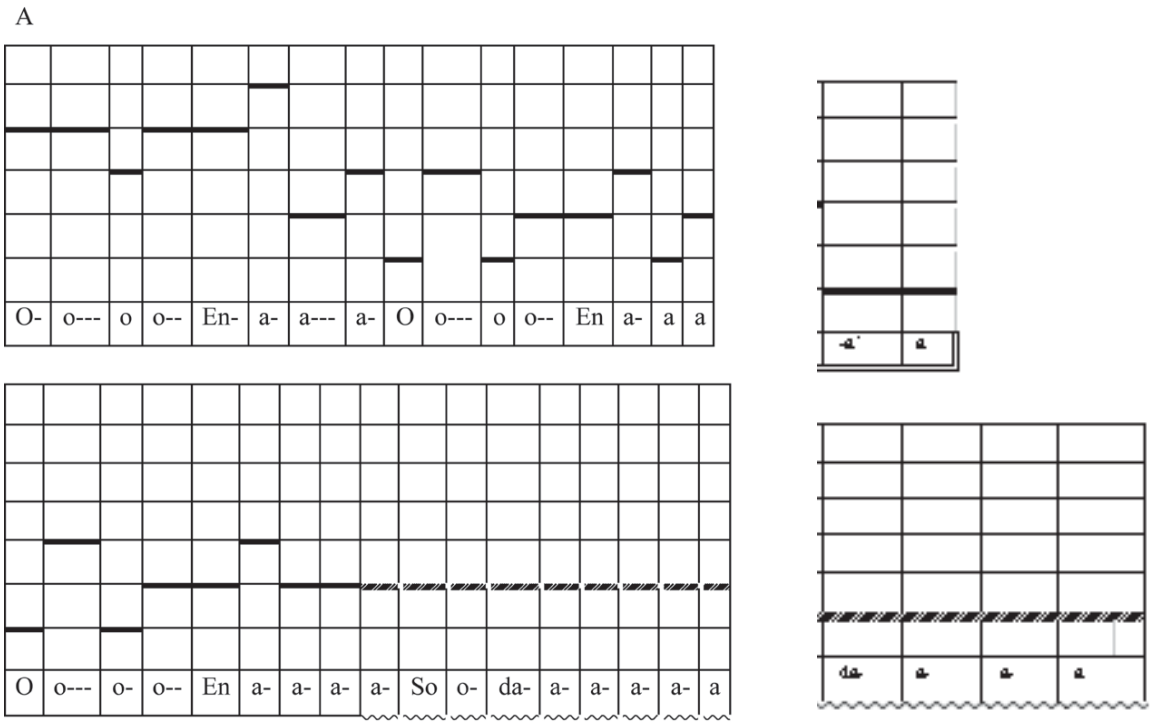


Figure 8: Mourning Song: *Dolo K'adi.* Stanza A.
Sound file PhA-JB2013_KLST_Fig8-9_dolo-k'adi.mp3

The line *Oh-oh-oh-oh En'a'a, oh-oh-oh-oh-oh En'a'a, oh-oh-oh-oh-oh En'a'a Soda* (Figure 8) appears three times in the performance with other stanzas in between as a kind of refrain. The overall form will be A B A C A. The A part has a binary form, a building block, descending motion and ends with a tone repetition that is sung as a *pulsating tone*, that is with a stress coinciding with each beat, which is a common trait in Native American music (denoted by a shaded line in the graph). This type of singing is

phonetically marked at the onset of each beat by a glottal constriction. The *dratax ch'elik* form is sung without a drum, unlike a dance song, and in the sections where a few syllables are stretched over a number of beats, the glottal pulses serve to keep the beat.

The stanzas B and C (Figure 9) consist of words with meaning and have no vocables, which is quite different from the dance songs. These stanzas consist of structured praise for the departed. Some *dratax ch'elik* have four or more verses; in this version of *Dolo K'adi*, two were recorded. Minto elders have suggested that a proper number of text verses to be used in such compositions is three: verses may focus on the feelings of the composer/ singer, on the important virtues of the departed, and lastly, on something that the departed person enjoyed doing. Examples of older songs recorded today may include fewer verses than were originally composed, because the designated song leaders cannot remember all the words.

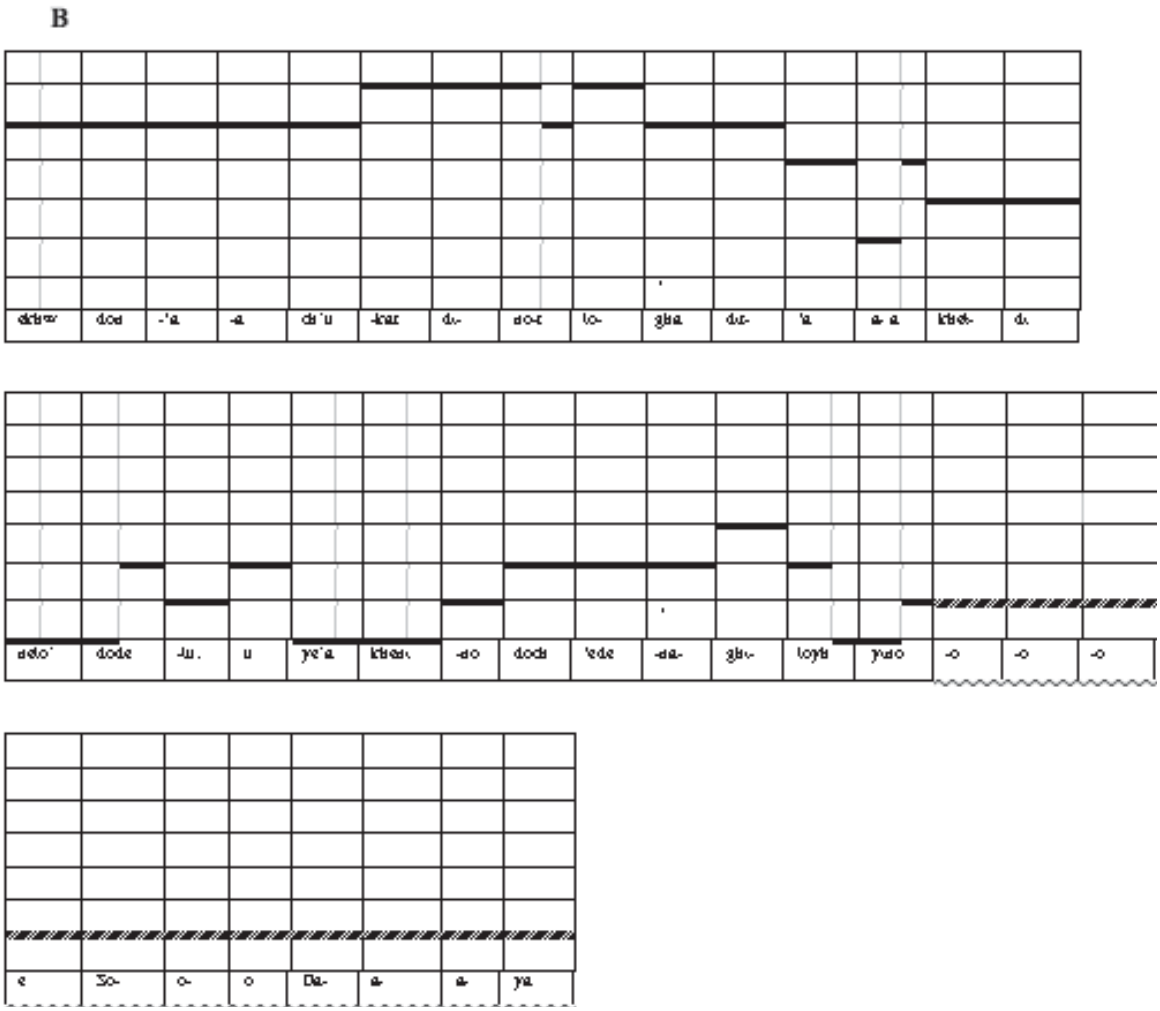


Figure 9: Mourning Song: *Dolo K'adi.* Stanzas B and C.
Sound file PhA-JB2013_KLST_Fig8-9_dolo-k'adi.mp3

Stanza B starts high and descends until finishing on a pulsating tone repetition. It is longer than A and contains more words. Basically this section can be seen as an extended variation of A that in the descending section has motifs that are slightly similar to the building block of A. It will take further research and comparisons between versions and with other examples of *dratax ch'elik* in order to establish whether this can be seen as a technique of composition.

The setting of the content words in *dratax ch'elik* differs from the text setting in dance songs, especially in the B and C sections, which have far more linguistic content than the A sections. In most cases low tone is not represented in the melody. However, morphologically prominent verb and noun stems have a tendency to be placed on downbeats, and are more likely to be lengthened or to fall on a moving note. This reflects the phonetic prominence of stem morphemes in speech, and probably also aids listeners in parsing.

3. *Silaa* (Venetie, Gwich'in/Koyukon). Sound file *3-silaa.wav.*

This song, which seems to belong to the 'sorry song' category, was recorded in Venetie by Silas John as part of a 'Songs and Legends' collection created during the 1970s. Mr John says in the interview surrounding the song that his father often sang it, and that he did so when he felt 'sorry'.

This song might be expected to be sung in Gwich'in, since Venetie is a Gwich'in village. However, not all words in the lyrics make sense in Gwich'in. The meaningful words in the song include *silaa,* which is clearly a kin term marked with a first person singular possessor. However, this is not a Gwich'in word. It is most similar to Koyukon *sel'aa* 'my uncle (mother's brother)'. The other words that seem meaningful are *sii dzaa niheii* 'come back here to me', which makes sense in Gwich'in. The rest of the words are interpreted by Gwich'in speakers as vocables.

(6) *Silaa*

A

1 Hileii heii ʾeii heii lii yooʾooho *silaa* haa haa haa
2 ʾEii heii ʾiya hoʾo ʾiya hah
3 ʾEii: *silaa* ʾohoʾ ahaha ʾa:ʾa

B

4 ʾEii: ʾiyo ʾolo *sii dzaa niheii*
5 Yatsii yaʾaa ʾeii: ʾiyoʾ hahaʾ
6 *ʾEna*ʾ, *silaa*, ʾahaʾ ʾeii heii ʾahahah

C

7 ’Eii: heii, ’eii heii ’iyo ’oho heii::::

8 ’Iyo’ oho’ ’iyeii, *silaa:::, ’aha’, ’aha’*

9 ’Ileii: hei’ ’iyo ’oho::: *’Aha’, silaa,* ha ha ha

Gwich’in musical practice shifted relatively early in the history of contact, moving to western forms of music such as European fiddle music, and singing with other instruments such as guitars. While this recording is relatively old compared to the Tanana recording, it still reflects the dependence of Gwich’in people on their more conservative Koyukon neighbors and relatives, who retain even now a lively tradition in memorial and dance songs. *Silaa* is more similar to Koyukon memorial songs, mourning songs converted to dance songs, than it is to the Tanana sorry song given above. It is rhythmic and upbeat, not lyrical like *Dolo K’adi.* Similar forms are seen in Johnston et al. (1978). Many Gwich’in families are connected to Koyukon families, so despite the sharp distinctions between these neighboring languages and traditions, social and cultural interdependence is also present.

A

1)

		▬	▬	▬	▬	▬	▬	▬	▬	▬			
											▬	▬	▬
▬	▬												
Hi	leii	heii	’eii	heii	lii	yoo	oo	ho	Si	laa	haa	haa	haa

2)

▬		▬	▬		▬			
	▬			▬				
						▬		▬
							▬	
’Eii	heii	’i	-ya	ho	-’o	’i	-ya	hah

3)

▬	▬	▬	▬	▬				▬	▬	
					▬	▬	▬			▬
’Eii	Si	-laa	’o	-ho’	a	-ha	-ha	’a	-a	’a’

Figure 10: *Silaa.* Sound file PhA-JB2013_KLST_Fig10_silaa.mp3

The song is strophic and though there are many similarities between the three stanzas there is also considerable variation both verbally and musically. It is not so easy to find a building block and the song is not isorhythmic. The general outline of the melody is falling: most of the 1st and 2nd line are in the higher area whereas the 3rd line is in the low area and the stanza ends with tone repetitions on the two lowest pitches. A binary form can be seen with a first section in the higher pitches, and a second section starting on the three last syllables of the 2nd line.

The high-pitched part has a very notable shift between two high tones that are very close to each other (roughly a semitone, illustrated in the graph (Figure 10) by the smaller distance between lines 2 and 3 counting from the top). When listening this can be experienced as a wavering or wailing pitch. This occurs in sad songs but seems not to be limited to this category. It can be seen as a downward bend on a tone, sometimes called a 'dip'. In many cases it is microtonal and has the character of an embellishment.

The words that carry meaning (*silaa* and *sii dzaa niheii* in the 2nd stanza) are at the end of the first line and highlighted by being sung as wavering high-pitched tones. Throughout the song the word *silaa* is, so to speak, embedded by vocables that are repeated throughout the stanza with variation. The basic structure of this song can be described as a binary form where the first section consists of a wavering at a high pitch and the second section consists of a finalizing tone repetition in the lower region, which seems to be the carrying idea.

4. Mixed lyrics: *Steamboat* (Venetie).
Sound file *5-venetie-steamboat-song.wav*.

"Steamboat" is a personal song, with no ceremonial function. Songs with lyrics in Athabascan and English became common in the 20th century, and are some of the most popular songs; their composers are often remembered. The Minto 'Airplane' song is a perennial favorite and has been popularized in a video; in this song, the airplane is a positive force, a ride to the potlatch.[4]

Another Minto dance song inserts English words such as 'whiskey' in the frame '_____, *enu'ey*!' ('Away with it!'). In that case, the influence of white outsiders is disapproved of. The Venetie song *Steamboat* paints a vigorous picture of the riverboat that used to run from Tanana to Fort Yukon, stopping in Stevens Village. The steamboat's status in the song resembles that of the Minto Airplane. Like *Silaa*, this song was recorded

4 See the YouTube video of this song in staged performance: "Minto Dancers Airplane Song and TailDuck song" (<http://www.youtube.com/watch?v=0aZybHWyCvE>).

by Silas John in 1972. In this lyric the words that have been translated are put in italics. English words and vocables are in plain print.

(7) *Steamboat*

Oni'ee steamboat	*dhee ni'ee*	Fort Yukon haa	*nijin na'oo*
3-back-comes	comes-back		wherever you2-go

'Steamboat's coming back to FortYukon – wherever you two go'

Saa	*dzaa*	*yuuzaa*
1SO-to	here	2s-come

'You come here to me'

'Ooni'eii hei 'ei down Steven Village 'ei hiyo
3-comes back

Tanana	*gwich'in naii* aa hee ee hee iya'oo
Tanana	person-PL

'Tanana people'

Goodbye steamboat	*oni'ee* yoo hoo ooo	*shalak naii*
	come-back	my-relative-PL

'Come back, my relatives'

Heii-eii-eii 'iya'a ha'a' 'ahahaha'aha

'Eii steamboat Fort Yukon *ooni'ee* hee ee haa aa ee ee hee 'iya haha
'Steamboat's coming back to Fort Yukon'

Oo Stevens Village	*ts'i*	*hi'iya'*	*'aha'*	*'aha'*
	from	they	are	going

'Leaving Stevens Village'

Tanana people 'eii, ha ha ha'

Sound file PhA-JB2013_KLST_ex7_steamboat.mp3

Songs like this one use English words because they deal with things that came with the English language. While the Athabascan languages typically borrow few words from languages with which they have contact, mixed songs deliberately foreground English language, sometimes to celebrate the things that came with it, sometimes to condemn them. In the *Steamboat*

song, the words *steamboat, Stevens Village* and *Tanana* are all foregrounded by the use of high pitch, extreme lengthening, and high intensity in the performance. The use of English to refer to Stevens Village and Tanana emphasizes the relationship between the places and the new, imported technology – the steamboat – that allows their rapid connection. The stress patterns of the English words are not disrupted in the text setting, but are expressed directly. Gwich'in speakers found this song easy to translate, suggesting that the Gwich'in words are also set directly to the rhythm of the song. Nevertheless, the beat of the song is never subordinated to speech-like constraints. This song contains much more actual text than the Minto *Airplane* and *Whisky* songs, both of which contain just one meaningful phrase, followed by rhythmic vocables.

Structurally this song has very much in common with the 'sorry song' above. It consists of one section, which is varied for the different text lines. Most of the first part of the section is in the high area making extensive use of dips in combination with pulsating tones. A main difference, though, is that this song is dominated by words with meaning whereas the vocables are few. The English loanwords are treated just like Athabascan words with dips and pulsation.

5. Medicine Song: Caribou Song (Tanana).
Sound file *5-venetie-steamboat-song.wav*.

This is a very old song that has been preserved in a teaching story. It is an 'animal song', but also counts as *senh ch'eliga'*, a medicine song. Such songs are said to have been taught to the people by the animals themselves. This song's melody recognizes the high-rising tone on the negative suffix, a very salient prosodic feature of the language. Songs of this type are probably among the oldest we have record of, and in former times they had ceremonial function. It should be noted that when listening to other old songs within stories, elders sometimes re-sang with different tunes.

The 'Caribou Song' was one of a set of medicine songs recorded by the highly respected elder Peter John with an unidentified interviewer.[5] This recording is not dated, but is likely to have been made in the 1970s. The story about it was recorded by the late Neal Charlie in 2008, and is a well-known teaching story that explains the lyrics. Other elders, such as Neal's brother Robert, also know and repeat the story and the song. The free sharing of this material contrasts with a general caution regarding medicine songs on the part of the remaining elders of Minto. Since the introduction of Christianity to the Tanana Valley by missionaries in the late

5 Alaska Native Language Archive, University of Alaska, Fairbanks, archive number TN27, ANLC2549

19th and early 20th centuries, expressions of Native Athabascan spirituality have been handled very discreetly in this area. Neal Charlie expressed his feelings about this music in the interview in which the story was recorded; he stated that medicine songs and other spiritual material could be misused if they were shared without full understanding of their proper application. For this reason, he was not willing to discuss many of the songs recorded and archived by Peter John. The Caribou Song was an exception because of its relationship to the teaching story, which is a lesson regarding truthfulness and obedience. Nevertheless, the form of this song resembles the form of Peter John's other examples in being more speech-like than musical.

(8) *Do* *seł̨dini* *chu*
what 1sO-3sS-TELL and
'What did you tell me?'
Bekwlá *seł̨dini* *chu*
3O-AR-BE-NEG 1sO-3sS-TELL and
'There are none, you tell me'

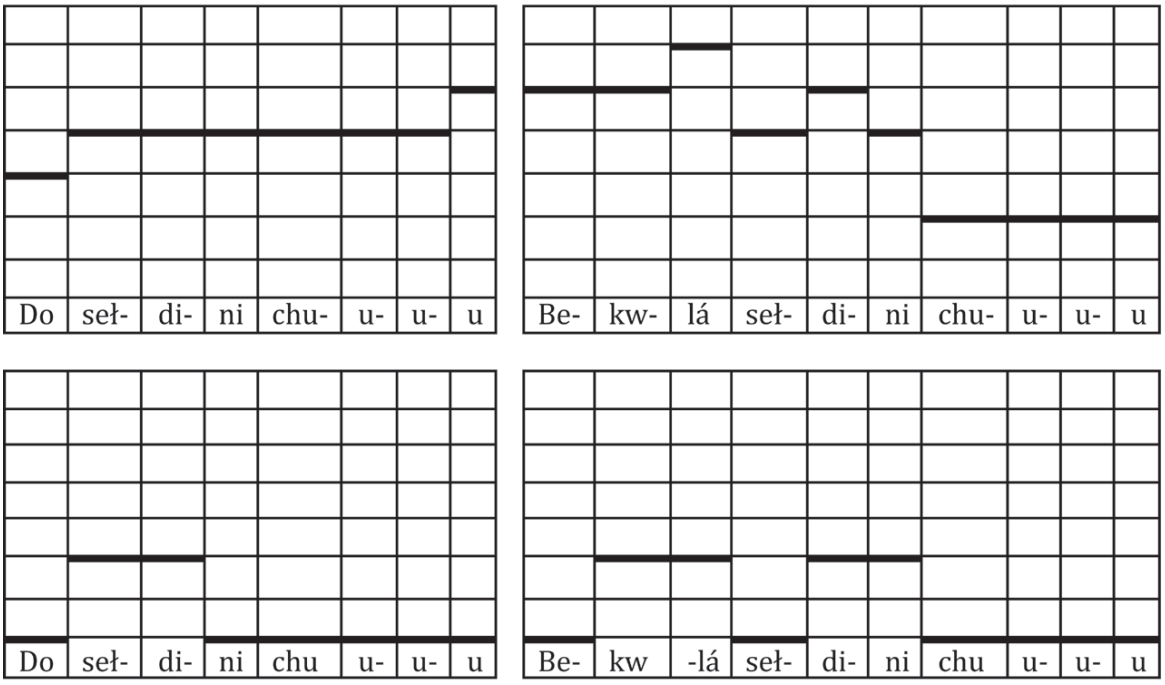


Figure 11: Caribou Song. Sound file PhA-JB2013_KLST_Fig11_caribou-song-neal.mp3

So-called animal songs are often short. This one consists of two sets of words making up a stanza. The binary form of this song is obvious since the text is sung twice. In the 1st section *Do seł̨dini chu* is given 8 beats

and *Bekwlá seł̨dini chu* 10 beats. This means that the two extra syllables of *Bekwlá* are simply added with one beat for each syllable. Apart from this the principle is isorhythmic and the first motif serves as a building block. The descending melody line is also obvious. There are no real tone repetitions to finish off the lines, but this can be explained by the fact that this is not a dance song. It should be noted that this song has no vocables without lexical meaning. Rather, it is one sentence repeated once isorhythmically in a descending fashion.

In contrast to the dance songs and mourning songs, the melody and rhythm of this song closely follow the rhythm and pitch contour of speech. The highest point in the melody falls on the negative stem in the word *bekwlá* 'there are none'. In the Minto dialect of Tanana, the negative and a few other specific words are marked with a high-rising, nasalized tone that is very distinctive in the language's prosody. In this word, the suffix is melded with the stem of the verb 'to be' to create a marked negative stem. The melodic setting recognizes the tone on this stem, which is the only marked syllable in the text. Rhythmically, the setting of one beat per syllable does not recognize the two light syllables in the text (*be-* and *kw)* as meriting different, lesser weight. Nor do the lexical stems in the text get special rhythmic treatment; *chu*, a conjunction, is elongated, but *–ni* 'say' gets just one beat, as does the high *-lá*. However, the overall contour, starting higher, finishing lower, with the negative emphasized, closely matches the contour of a Minto intonational unit. We take this to reflect the importance of the lyrics in this song.

Concluding discussion

There is clear evidence that the two lexical tones in the tonal Yùan Kammu are present in the vocal genres, though combined with musical pitch in different ways in specific vocal genres, spanning from close to 100% correlation to around 50%. The lower the correlation, the more dominant is the melodic template the vocal genre is built on. In such vocal genres some parts of the performance – normally the beginning and end of musical phrases – are characterized by music pitch centration and the other part as lexical tone centration.

In the Athabascan material there are songs that recognize the high-rising tone on the negative suffix, a very salient prosodic feature of the language. In contrast to the dance songs and mourning songs, the melody and rhythm of such songs closely follow the rhythm and pitch contour of speech. The highest point in the melody falls on the negative stem in the word *bekwlá* 'there are none'. In the Minto dialect of Tanana, the negative and a few other specific words are marked with a high-rising, nasalized

tone that is very distinctive in the language's prosody. The distinction music pitch centration / lexical tone centration seems to apply also here.

Rhythmically, the setting of one beat per syllable in an Athabascan medicine song does not recognize the two light syllables in the text as meriting different, lesser weight. Nor do the lexical stems in the text get special rhythmic treatment. However, the overall contour, starting higher, finishing lower, with the negative emphasized, closely matches the contour of a Minto intonational unit. Similar practices were noted in the Kammu tradition.

The occurrence of lexical tones in tonal Kammu suppresses the use of sentence intonation. In singing (Guiding song and other genres), tonal Kammu uses three pitch levels for the realization of whole phrases. This can be seen as a way to preserve lexical tones in singing but may be done for other reasons in spoken Guiding songs. How this works in non-tonal Kammu singing needs to be investigated in more detail. Within the area of Athabascan song, some song types map melody more closely to speech than others, recognizing syllable weight and lexical tone.

Accepting that speech also has its rhythm we can observe how rhythmization is achieved in the transition between Kammu vocal genres. Thus, rhythm in Guiding songs builds on the increased degree of speech rhythm. Storytelling uses aesthetic means not occurring in everyday speech but still not conflicting with the phonological system of the language.

Possibly, when prosody is taken into account, the distinction music pitch centration / lexical tone centration should be seen as a special case of a more general distinction: music centration / language centration.

A comparison between speech and song may lead to a more detailed analysis of the speech prosody of a language. When we observe differences between tonal and non-tonal speakers in using pitch registers in songs, this should also be analyzed in speech. Also, the proposed comparison may be a challenge for existing prosodic models. As we may find speech features that are more robustly preserved than others in song we may suppose that they should be captured by prosodic models.

Concerning the idea of a continuum from speech to song a prosodic continuum could not be established for Kammu as we did not find a continuous transition between vocal genres. Instead every genre picks up some prosodic feature to create rhythmicity and to contrast with everyday speech. A possible explanation is that the different vocal genres also have distinct functions in specific contexts, which means that they must be different enough to serve as markers for these functions and contexts. Therefore the continuum would be one of steps rather than gradual continuity.

The suggested continuum speech to song in Kammu culture was built on musical and poetic criteria. If we take speech as a starting point for comparison we find the following transitions: everyday speech > everyday narratives > spoken Guiding songs > sung Guiding song; everyday speech > story-telling; everyday speech > *hrlɨ̀ɨ̀*. This result implies that – if it is desirable to define such a continuum – it will have to be thought of as a three-dimensional continuum.

Athabascan musical metavocabulary distinguishes clearly between speech (*kena* 'word') and song (*ch'elik* 'song'). The functions of spoken and sung genres are clearly defined. In the examples we have shown here, a medicine song within a story shows the closest relationship to speech; next most speech-like are the content verses within a memorial song. A mixed-language song seems intermediate between the memorial song style and the dance song style, in which meaningful words are minimized and rhythm and melody predominate.

It is yet too early to tell whether the existence of a continuum from speech to song is culture specific to Kammu and related cultures or if also the study of Athabascan vocal expressions can gain from arranging them as a continuum.

The study of Kammu songs has shown that performance can be understood as the re-creation of poetic and musical material organized according to specific templates. If similar templates may be identified as a method of analysis of Athabascan songs these templates are less obvious and differently organized.

The Athabascan songs studied here have different structures. A song may be built on a poignant melodic/rhythmic setting of a verbal phrase serving as a building block which – combined with vocables – is sequenced in downward motion and ends with a tone repetition on the tonic. In another case a binary form where the first section consists of a wavering at a high pitch and the second section consists of a finalizing tone repetition in the lower region seems to be the carrying idea. This suggests that it is worth to pursue this line of analysis further and that the combined linguistic and ethnomusicological approach may prove more useful than ordinary musical analysis.

The work within our research project will focus on developing a method of analysis of vocal expressions involving the perspectives of ethnomusicology, prosody, syntax and semantics. In this joint study we have made a first attempt to put some of these approaches together. Each approach brings with it specific knowledge, methodologies and terminology. The process to integrate – to a certain degree – these approaches has just started and will be a central matter in our continued work.

This study has shown that methods developed in the study of Kammu culture only partly apply to Athabascan culture and vice versa. This is true for the construction of vocal expressions: the Kammu genre-specific melodic templates and the Athabascan building blocks. It has also shown that methods developed in ethnomusicology do not always correlate with those developed within linguistics: the continuum from speech to song, based on vocal genres and poetics in Kammu culture, does not easily apply for prosody.

Our methods continue to develop in our search for a methodology for the study of speech and song that is not culture-specific. More research will be needed that will include the study of existing data from new perspectives, the search for new data from these cultures and also data from other cultures.

Please visit http://www.phonogrammarchiv.at/7DE709JKSAE
for additional appendices.

References

Asano, Yuki & Bettina Braun. 2011. "Integrating lexical and postlexical suprasegmental information in native and non-native Japanese". Paper presented at the NINJAL International Conference on Phonetics and Phonology (ICPP 2011), Kyoto (Japan).

Berez, Andrea L. 2011. "Prosody as a genre-distinguishing feature in Ahtna: a quantitative approach". *Functions of Language* 18 (2): 210–237.

Coray, Craig. 2007. *Dnaghelt'ana Qut'ana K'eli Ahdelyax. (They Sing the Songs of Many Peoples). The 1954 Nondalton recordings of John Coray.* Nondalton: Kijik Corporation.

Holton, Gary. 2005. "Pitch, tone and intonation in Tanacross". In: Rice, Keren & Sharon Hargus (eds.). *Athabaskan prosody*. Amsterdam: John Benjamins, 249–276.

Honigmann, John J. 1981. "Expressive aspects of Subarctic Indian culture". In: Helm, June (ed.). *Handbook of North American Indians.* Washington: Smithsonian Institution, 718–738.

Hymes, Dell. 1981. *"In vain I tried to tell you": Essays in Native American Ethnopoetics.* Philadelphia: University of Pennsylvania Press.

Johnston, Thomas F. 1994. "The social role of Alaskan Athabascan Potlatch dancing". In: Overby, Lynnette & James H. Humphrey (eds.). *Dance: current selected research*. Vol. 3. New York: AMS Press, 183–226.

Johnston, Thomas F. et al. (eds.). 1978. *Koyukon Athabascan dance songs.* Anchorage: National Bilingual Materials Development Center.

Karlsson, Anastasia, David House & Jan-Olof Svantesson. 2012. "Intonation adapts to lexical tone: the case of Kammu". *Phonetica* 69 (1–2): 28–47.

Krauss, Michael E. 2005 [1979]. "Athabaskan Tone". In Hargus, Sharon & Keren Rice (eds.). *Athabaskan Prosody.* Amsterdam: John Benjamins, 51–136.

List, George. 1961. "Speech melody and song melody in central Thailand". *Ethnomusicology* 5 (1): 24–26.

List, George. 1963. "The boundaries of speech and song". *Ethnomusicology* 7 (1): 1–16.

Lord, Albert B. 1960. *The singer of tales*. (Harvard Studies in Comparative Literature, 24). Cambridge: Harvard University Press.

Lovick, Olga & Siri Tuttle. 2012. "Prosody of Dena'ina Narrative Discourse". *International Journal of American Linguistics* 78 (3): 293–334.

Lundström, Håkan. 1980. "North Athabascan Story Songs and Dance Songs". In: Rooth, Anna Birgitta (ed.). *The Alaska seminar*. Stockholm: Almqvist & Wiksell, 126–164.

Lundström, Håkan. 2010. *I will send my song. Kammu vocal genres in the singing of Kam Raw.* Copenhagen: NIAS Press [with CD].

Lundström, Håkan & Jan-Olof Svantesson. 2008. "Hŕlìì singing and word-tones in Kammu". *Working papers* (Dept. of Linguistics and Phonetics, Lund University) 53: 117–131.

Mithun, Marianne. 1999. *The languages of Native North America.* Cambridge: Cambridge University Press.

Patel, Aniruddh. 2008. *Music, language, and the brain.* Oxford: Oxford University Press.

Pearce, Tony Scott. 1985. *Musical characteristics: Tanana Athabascan dance songs.* MA thesis, University of Alaska Fairbanks [with sound recordings].

Powers, William K. 1992. “Translating the Untranslatable: The Place of the Vocable in Lakota Song”. In Swann, Brian (ed.). *On the Translation of Native American Literatures.* Washington: Smithsonian Institution Press, 293–310.

Seeger, Anthony. 1987. *Why Suyá sing. A musical anthropology of an Amazonian people.* Cambridge: Cambridge University Press.

Shaul, David L. 1992. “A Hopi Song-Poem in ‘Context’”. In Swann, Brian (ed.). *On the Translation of Native American Literatures.* Washington: Smithsonian Institution Press, 228–241.

Slobodin, Richard. 1981. “Kutchin”. In: Helm, June (ed.). *Handbook of North American Indians*. Washington: Smithsonian Institution, 514–532.

Svantesson, Jan-Olof. 1983. *Kammu phonology and morphology.* Lund: Gleerup.

Svantesson, Jan-Olof. 2004. “Syllable boundaries in Kammu”. In: Branderud, Peter & Hartmut Traunmüller (eds.). *Proceedings, Fonetik 2004*. Stockholm: Stockholm University, 152–154.

Svantesson, Jan-Olof & David House. 2006. “Tone production, tone perception and Kammu tonogenesis”. *Phonology* 23: 309–333.

Svantesson, Jan-Olof & Anastasia Karlsson. 2004. “Minor syllable tones in Kammu”. In: Bel, Bernard & Isabelle Marlien (eds.). *Proceedings of the International symposium on tonal aspects of languages* (*TAL2004*). Beijing: Chinese Academy of Social Sciences, 177–180.

Tayanin, Damrong (Kam Raw). 2006. *Where do our souls go after death? Kammu traditions, rituals and ceremonies.* Bangkok: s.n.

Tuttle, Siri G. 1998. *Metrical and tonal structure in Tanana Athabaskan.* PhD thesis, University of Washington.

Tuttle, Siri G. 2003. “Archival phonetics: stress and tone in Tanana Athabaskan”. *Anthropological Linguistics* 45 (3): 316–336.

Tuttle, Siri G. 2011. “Language and music in the songs of Minto, Alaska’. *Language Documentation and Description* 10: 82–112.

Tuttle, Siri G. 2012. “The margins of tonality: a case from Alaskan Athabaskan”. Paper presented at the workshop “Tone: Theory and Practice” at the Max Planck Institute for Evolutionary Anthropology, September 28–29, 2012, Leipzig (Germany).

Melodic Relativisation of Speech Tones in Classical Vietnamese Singing: the Case of *Many Voices*

(Gisa Jähnichen)

Introduction: *Many Voices* (*36 giọng*)

This paper analyses the strategies in relativisation of speech tones in classical Vietnamese singing. The piece *36 giọng* (*Many Voices*[1]) is an additional part of the classical *hát ả đào*, also called *ca trù*. This North Vietnamese *ca trù* has recently been inscribed on the World's Intangible Heritage list as being in urgent need of safeguarding.[2] Its roots reach back to the 11th century (Jähnichen 1997, 1: 3–9; Đỗ Bằng Đoàn & Đỗ Trọng Huề 1962). Present *ca trù* sessions are performed by a female vocalist who also plays a bamboo slab, one lute player, and mostly a drummer who is considered to be part of the audience that elects him due to his profound knowledge of poetry and music.

Within the repertoire of *ca trù*, *Many Voices* belongs to the last pieces of the entertainment section usually played shortly before the performance session ends (appendix 1). It consists of 17 different sections quoted from various sources such as excerpts from the main *ca trù* repertoire, the rural *chèo* theatre and its folk song adaptations, and the music of temple ceremonies. The 17 different sections can be grouped into free metric and solid metric sections, in sections that consist of parts with lyrics related to each other, and in sections that follow a similar musical mode.

In previously published local and academic literature (Ngô, Linh Ngọc & Văn Phú Ngô 1987: 238ff.; Trần, Văn Khê 1982: 168f.), the following assignments of different parts in *Many Voices* can be found:

1 The direct translation is "36 voices", 36 being a metaphor for "many".

2 *Ca trù* of the Việt people was included in the respective UNESCO list on 1 October 2009, at the fourth session of the Intergovernmental Committee for the Safeguarding of Intangible Cultural Heritage in Abu Dhabi.

			Assignment according to:			
			material		context	
Section	No.	"part"	literary	musical	social	functional
1-2	1	sa mạc	recitation of verse poetry			
	2	bổng mạc	recitation of verse poetry	with shifting reference tone		
3	3	xướng tế	recitation of ritual text in the manner of novices			announcing the order of ceremonial acts
4-6	4	đò đưa		songs	of the boatmen	
	5	huê tình		song	love	
	6	trống quân		call and response singing	of young ladies and men on the countryside	
7-8	7	nói sử	recitation			of chèo theatre poetry in the serious, so called 'historical' style
	8	bài sai lên đồng		song	of the posessed	
9-13	9	kể truyện sa mạc	narration of stories	in the style of sa mạc		
	10	thổng thiên thai		transition to the style thiên thai		
	11	hát cách		chant		of the chèo theatre repertoire
	12	chầu văn			shamanic	incantation
	13	hãm		song	drinking	
14-17	14	sa mạc tỳ bà	performing of two lines	from the *Long Song* (tỳ bà hành) in the style sa mạc		
	15	tỳ bà hành	Two lines of the poem tỳ bà hành			
	16	cung bậc		singing in three different modes		
	17	dựng tỳ bà cung huỳnh		musical mode cung huỳnh from the *Long Song* tỳ bà hành		

Table 1: Assignments of different parts in *Many Voices* according to Ngô Linh Ngọc & Ngô Văn Phú (1987: 238ff.) and Trần Văn Khê (1982: 168f.).

All parts of *Many Voices* represent a literary as well as a musical abstract of their source repertoire. This abstraction process is a result of genre establishments during early urbanisation of music practices developed from distinct rural traditions (Jähnichen 1997, 1: 78ff.). The great diversity of characteristics assigned by earlier scholars shows further that poetry and musical ideas are explained from various perspectives depending on available attributes handed down over the centuries.

In this study, I take one complete sound recording as an analytical support (appendix 4). This recording was made in 1993 in the Daoist temple *đền dâu* of Hanoi's old city centre. It features the artists Bạch Vân (voice and bamboo slab, called *phách*), Chu Văn Du (lute, called *đàn đáy),* Trúc Hiền (drum, called *trống chầu*), and Phó Thị Kim Đức, who assists the drummer on a further small drum usually applied in excerpts deriving from the *chèo* theatre repertoire. However, drummer, assistant drummer and the bamboo slab do not play in a classical *ca trù* style. They add some 'colour' to the general mood of the performance and are not considered as being essential in *Many Voices* (Ngô Linh Ngọc & Ngô Văn Phú 1987: 239).

Other recordings of that piece confirm that drum and bamboo slab can be omitted for most of the duration of the piece (Jähnichen 1997, 2: 37–48).[3] Therefore, I leave bamboo slab and drums out of the analysis although they play a role for the creativity of the performers. Leaving this observation to further investigations, this study focuses on the relationship of language characteristics and melodic lines in the music performance of a piece that abstracts and summarises historically and functionally different layers of lyrics that might be represented in related musical thoughts.

Lyrics and Song Texts

The pivotal lyrics used in *Many Voices* are as follows and are here translated for understanding the diversity in the meaning of different sections:

part	pivotal lyrics[4]	translation (in lines)
1	tôi chi anh Hoe tôi chi anh Tuyết, tôi chi anh Nguyệt nay anh Phong Hoe mai anh Tuyết Nguyệt nay anh lấn lận, mai anh lựa lựa lấn lận, lựa lựa	I with Hoe[5], I with Tuyết, I with Nguyệt today Phong Hoe, tomorrow Tuyết Nguyệt, today you are unfaithful tomorrow you are choosy;
2	còn gì là xuân bật[6] phong trần phải phong trần cho thanh cao mới được phần thanh cao	what remains of spring time; who has a hard life must learn to cope with it, one's fate decides on a noble-minded life;
3	mặc áo the thâm đứng dựa cột đình xướng rằng: cù soát tế vật tự lập cự ai!	in a dark robe leaning against a pillar of the communal house and reciting loudly: all have to pay tribute, though being against somebody!
4	chân bước xuống thuyền thời thuyền chả giậm chả đi giậm ra ván nát thuyền thì long đanh	entering a boat, without punting, the boat does not move, but the rowing rhythm causes the planks to burst, the swinging boat is getting nowhere;
5	công anh lên thác xuống ghềnh lên thác đã vậy xuống ghềnh làm sao...	you went the waterfall up and the rapids down, you did it. Why did you?

3 One recording made in 1995 features the singer Bà Mùi, one of the most outstanding performers of *ca trù* until the late 1990s, who is just accompanied by the lutenist Phó Đình Kỳ.

4 The lyrics are set in traditional line scheme according to the poetic line order, incomplete lines are indented.

5 Hoe is actually the name 'Huê'. The pivotal lyrics of the piece are, nevertheless, transcribed unchanged from generation to generation. (The clarifying footnotes 5 to 9 were inserted at the request of the anonymous reviewer, to whom I am also grateful for pointing out a few typographical errors in tables 1, 2 and 4).

6 Here, the word bật appears instead of the more common bậc. Similar cases of substituting final <c> with <t> can be found with bút rứt, which – in spoken language – might be bức rức.

6	ai làm đất thấp trời cao ngọn đèn sáng tỏ hơn sao đêm rằm Mừng đàn anh lại mừng dây mừng chim loan phượng ngô đồng tốt tươi	the earth is small, the heaven is big, the light of the lantern is brighter than the stars by night in the middle of the lunar month, rejoicing instruments and strings, rejoicing the phoenix on the oleaginous fruit tree (with nicely fresh fruits)
7	dạ thưa bác, bác giai có nhà hay đi vắng thưa với bác, thầy cháu lên tỉnh vắng chưa về mời bác ngồi chơi, tôi mời bác ngồi chơi, để tôi bảo trẻ quạt nước bác xơi mai sẽ về	Dear Lady, is my master at home or not? Dear Mister, my husband went into town and has not got home yet, please, sit down here and stay, I call the children to serve you tea, tomorrow you may return;
8	thôi thôi bà để tôi về mặt trời đã xế dặm hòe còn xa	No, please, let me go home, the sun is setting down soon, the way is still long;
9	Một vùng cỏ ấy bóng tà Gió hiu, hiu thổi một và[7] bông lau rút trâm sẵn giật mái đầu, vạch ra cây vịnh bốn câu ba vần	withered meadow, long shadows, the wind touches quietly some reed tips, pulling out some hair needles to engrave poetry of four phrases and three rhymes into the bark of a tree;
10	cỏ cây xem vẫn tần ngần lối vào chẳng biết rằng gần hay xa xinh thay hơi thú yên hà đào nguyên muốn hỏi ai là chủ nhân	still curiously looking at grasses and trees, not knowing whether the way is long or short, how wonderful, how pleasant is the beauty of nature... the fairy caves have yet to be asked for their man of the house;
11	Rất mực hồng quần xuân xanh xấp xỉ tới tuần cập kê	wearing pink pants, nearly in the age of becoming mature until the moment of getting married;
12	ngàn bà ngàn mái ngàn me ngàn giang, ngàn trúc, ngàn tre, ngàn vầu, dăm ba thiếu nữ theo hầu cô xe chỉ thậm, cô xâu hạt vàng	Out of thousand women, thousand roofs, thousand tamarinds, thousand kinds of long node bamboo, thousand kinds of short node bamboo, thousand bamboo reeds, thousand thick bamboos, there are only few young ladies who will do their duty; one lady is sewing nicely, another lady is threading golden seeds;
13	từ khi thiếp bén duyên chàng bướm ong sum họp phượng hoàng no đôi góc bể bên giời	since I am aware of your love, butterflies and bees are gathering in happy pairs around the golden phoenix at a quiet place off the sea in heaven;
14	vặn đàn vài tiếng dạo qua tuy chưa trọn khúc tình đà thoảng hay	Tuning the instrument and preluding on some strings, though the piece is not yet complete, a warm feeling floods through the mind;
15	nghe não nuột[8] mấy dây bứt rứt nhường than niềm tấm tực bấy lâu	If sad, sit down and listen to the sound of the strings that suppresses for some time sighs and anger;

7 Here, the word vài would fit better, yet for the rhyme with tà in the previous line the word has been transformed into và.

8 The word nuột is an old version of ruột in the dialect spoken throughout the Red River delta (in some areas also luột).

16	ngán thay cho vợ chồng Ngâu cách một dịp dàu[9] dạ nọ kém tươi	Bad things happened to the couple Ngâu a painful distance apart from each other and not young anymore;
17	lệ ai chan chưa hơn người Giang Châu Tư Mã đượm mùi áo xanh.	weeping countless tears, somewhere dashing around countries far away, feeling the smell of long roads in one's robe.

Table 2: Pivotal lyrics (translation by the author).

The eclectic combination of text excerpts from oral tradition as well as from classical verse poetry, especially the *Story of Kiều* by Nguyễn Du, indicates the familiarity of the audience with the literary sources.

To become a song text, the lyrics have to be transformed. In sung poetry, the six speech tones are categorised into high, low, and moving pitches (= contour tones):

model syllable	melodic contour	in poetry considered as	represented by analytical sign
ma		high (unchanging)	-
mà		low (unchanging)	_
mã		moving pitches	~
má			
mả			
mạ			

Table 3: North Vietnamese speech tones and their consideration in verse poetry (Jähnichen 1997, 1: 135–144; Brunelle 2009).

The melodic structure is supposed to behave accordingly, thus allowing only certain speech tone combinations to fit into given melodic patterns, though not in a very strict way (compare Stock 1999: 191). To avoid misunderstandings, auxiliary syllables, repetitions of text phrases, additional text phrases, and vowel extensions are used to relativise speech tones. This process of transformation is of varying sophistication in different Vietnamese vocal genres. The following comparison between pivotal lyrics and the song lyrics actually produced during the performance shows the resulting density of verbal content in relation to the, so to say, "functional" content of the lyrics.

9 Here, the word dù would be clearer to understand; however, for the rhyme with the word Ngâu in the previous line, the word has been transformed into dàu.

	pivotal lyrics	song text (**pivotal** / repetitions and additions / *auxiliary syllables*)
1	tôi chi anh Hoe tôi chi anh Tuyết, tôi chi anh Nguyệt nay anh Phong Hoe mai anh Tuyết Nguyệt nay anh lấn lận, mai anh lựa lựa lấn lận, lựa lựa	*Ư* anh Phong **tôi chi** *ư...* **anh Hoe** **Tôi chi anh Tuyết,** **tôi chi anh Nguyệt** *ư...* Tôi chi **nay** ư.... anh, **anh Phong Hoe** **mai anh Tuyết Nguyệt**... **nay anh lấn lận**, **mai anh lựa lựa**... **lấn lận, lựa lựa**
2	còn gì là xuân bật phong trần phải phong trần cho thanh cao mới được phần thanh cao	*ư...* còn... *ư...* gì... *ư...* **còn gì là** *ư... ư...* **xuân**... **bật phong... trần** thời **phải**... **phong... trần** **cho thanh cao mới... được... phần**,... **thanh** *n...* **cao**... *hư...*
3	mặc áo the thâm đứng dựa cột đình xướng rằng: cù soát tế vật tự lập cự ai!	**mặc áo the thâm** *m...* **đứng dựa cột**... **đình** **xướng rằng: cù**... **soát tế vật** *n...*, **tự**... **lập** *n...*, **cự**... **ai!** *Hơ...* anh *ơi*.....
4	chân bước xuống thuyền thời thuyền chả giậm chả đi giậm ra ván nát thuyền thì long đanh	**chân** *i* **bước xuống thuyền**, thuyền **chả** có **giậm**, **thời** *i* **thuyền chả** có **đi** *ng...* **giậm ra** là ra **ván nát** *i...* chự thuyền thì **thuyền thì** đanh chự **long đanh** *n...*
5	công anh lên thác xuống ghềnh lên thác đã vậy xuống ghềnh làm sao...	Tối biết **công anh lên thác** *i* **xuống ghềnh** *ng...* Lên thác ai *ơi* xuống ghềnh **lên thác** rồi ra **đã**... **vậy** *ấy mấy* xuống... ghềnh *ư...* **xuống ghềnh** sao... chự... **làm sao**...
6	ai làm đất thấp trời cao ngọn đèn sáng tỏ hơn sao đêm rằm Mừng đàn anh lại mừng dây mừng chim loan phượng ngô đồng tốt tươi	Đất thấp ông trời... cao **ai... làm** thời **đất thấp** ông **trời**... **cao**... **ngọn... đèn** thời **sáng tỏ hơn sao đêm... rằm**... **Mừng đàn** thời **anh lại mừng dây** **mừng...chim** thời **loan phượng** ấy **ngô đồng** mà **tốt** *n...***tươi** *n...*
7	dạ thưa bác, bác giai có nhà hay đi vắng thưa với bác, thầy cháu lên tỉnh vắng chưa về mời bác ngồi chơi, tôi mời bác ngồi chơi, để tôi bảo trẻ quạt nước bác xơi mai s¨ về	**Dạ thưa** với **bác**, **bác giai** tôi **có nhà hay đi vắng**... thưa với bác, **thầy cháu**... **lên tỉnh vắng chưa về** **mời bác ngồi chơi,** tôi mời bác ngồi chơi, **để tôi bảo trẻ quạt nước bác xơi... mai sẽ về**
8	thôi thôi bà để tôi về mặt trời đã xế dặm hòe còn xa	Bà để tôi... về **thôi thôi bà để tôi**... **về** **Mặt trời đã**... **xế**... **dặm**... **hòe còn**... **xa**...
9	Một vùng cỏ ấy bóng tà Gió hiu, hiu thổi một và bông lau rút trâm sẵn giật mái đầu, vạch ra cây vịnh bốn câu ba vần	**Một vùng cỏ ấy bóng tà**... **Gió hiu hiu thổi**...**một**...**và bông**...**lau**... **rút trâm sẵn**... **giật mái đầu**, **vạch ra cây**... **vịnh**... **bốn câu**... **ba**... **vần**
10	cỏ cây xem vẫn tần ngần lối vào chững biết rằng gần hay xa xinh thay hơi thú yên hà đào nguyên muốn hỏi ai là chủ nhân	Tần ngần, **cỏ cây xem vẫn tần ngần** **lối vào chững biết** *n...* **rằng**... *ư hư...* **gần**... **hay xa** yên hà **xinh**... **thay hợi**... **thú** *ng...* **yên hà** **đào nguyên muốn hỏi**... **ai** *ư hư... ư* **là** ư **chủ**... **nhân**.
11	Rất mực hồng quần xuân xanh xấp xỉ tới tuần cập kê	**Rất mực** *ấy mấy* **hồng** i **quần** i **xuân** *i* xanh *n...* là **xanh xấp** *i...* **xỉ** chự đã **tới** *ng...* **tuần** kê chự **cập** *i* **kê** *i*

12	ngàn bà ngàn mái ngàn me ngàn giang, ngàn trúc, ngàn tre, ngàn vầu, dăm ba thiếu nữ theo hầu cô xe chỉ thậm, cô xâu hạt vàng	Nói *i i* trên ngàn thời ... ngàn mái ngàn me... trên **ngàn bà ngàn mái ngàn me**... **ngàn giang, ngàn trúc** *n..*, **ngàn** me... **tre**, **ngàn vầu**, thiếu nữ theo... hầu, **dăm ba** cô **thiếu nữ** **theo**... **hầu** **cô xe** là xe **chỉ**... **thậm**, **cô xâu** *ư...*, xâu **hạt vàng**, *ng...* .
13	từ khi thiếp bén duyên chàng bướm ong sum họp phượng hoàng no đôi góc bể bên giời	Thiếp bén *i* duyên *i* chàng... *ng...* **từ khi thiếp bén** *i...* duyên *i* **duyên chàng** *ng...* **bướm**... **ong**... **sum họp** *n...* **phượng** *ư ư... ư....* *ư...* **hoàng**..., phượng... hoàng *ư* **no** ... **đôi** **góc bể** *i* **bên** *i* **giời**...
14	vặn đàn vài tiếng dạo qua tuy chưa trọn khúc tình đà thoảng hay	*Hư* **vặn** *n...* **đàn**... **vài tiếng**, tiếng **dạo**... **qua**... *ư... ư...* **tuy** *ng* **chưa** *ng* **trọn** *n...* **khúc** *ng* tình... *ư...* đà *ng...*, **tình đà thoảng**... *ư...* **hay**
15	nghe não nuột mấy dây bứt rứt nhường than niềm tấm tực bấy lâu	Bựt rựt **nghe não nuột** *n...* **mấy** *ư* **dây**... **bựt rựt** *n... n...* **Nhường than niềm tấm**... **tực**... *ư... ư...* **bấy** *ng* **lâu**
16	ngán thay cho vợ chồng Ngâu cách một dịp dàu dạ nọ kém tươi	Vợ...chồng, vợ chồng Ngâu... **ngán thay cho** n²i...vợ *ng...* chồng **vợ chồng** **Ngâu**... **cách** *n* **một** *n* **dịp**... **dàu dạ nọ** *ư...* *hư... ư...* **kém**... **tươi**...
17	lệ ai chan chựa hơn người Giang Châu Tư Mã đượm mùi áo xanh.	Hơn người **lệ ai** *ng* **chan chựa hơn người** **Giang Châu Tư Mã**... *ư* **đượm** *ư ư... ư...* **mùi** *ư* **áo**... **xanh**.

Table: 4: Lyrics and song text as recorded (see appendix 4).

A step by step comparison reveals that there are no strict rules transforming properties of the sources, such as the literary metre, the musical metre or reference tone, into the performance. One might think that free metric recitation would more accurately follow speech tones (see grey areas in the following table, also compare Clayton 1996: 324). This assumption may only apply in some cases, such as the first two sections as well as in section 9 and 10. All other musically free metric parts use nearly the same frequent amount of song text structuring elements. In the following table, the mentioned parameters are listed to allow comparison with the previous table. The sources used are actual performances, being elaborated through playing practice. Here, the source term *hát ả đào* is given instead of *ca trù*, which was a denomination established in the later period of *hát ả đào*. As further sources, the music of temple ceremonies *chầu văn*, local folk songs *dân ca* and *hát hội* in their festive appearance, and finally *hát chèo* are mentioned, the prevalent rural music theatre.

	part	source	literary metre	musical metre	possible style of phách playing	reference tone
1	sa mạc	dân ca	4 syllables (bốn chữ)	free	-	B
2	bổng mạc	dân ca	6-8 syllables (lục bát)	free	-	F#
3	xướng tế	chầu văn	free metre (văn tế)	free	-	F#
4	đò đưa	hát chèo	4+lục bát	solid	chèo style	B
5	huê tình	dân ca hát chèo hát à đào	lục bát	solid	chèo style	F#
6	trống quân	dân ca hát hội	lục bát	solid	chèo style	F#
7	nói sử	hát chèo	prose	free	-	B
8	bài sai lên đồng	chầu văn	lục bát (Story of Kiều)	solid - free	chèo style	B
9	kể truyện sa mạc	hát à đào	lục bát (Story of Kiều)	free	-	F#
10	thống thiên thai	hát à đào	lục bát (Story of Kiều)	free	khổ giữa	B
11	hát cách	hát chèo	4+ lục bát (Story of Kiều)	solid	chèo style	B
12	chầu văn	chầu văn	lục bát	solid - free	chèo style	B
13	hãm	hát chèo hát à đào	lục bát	solid	chèo style	B
14	sa mạc tỳ bà	hát à đào	lục bát	free	-	F#
15	tỳ bà hành	hát à đào	7 syllables (thất ngôn)	free	khổ giữa	B
16	cung bậc	hát à đào	lục bát (oral poetry with special rhyme)	free	khổ giữa	B / F#
17	dựng tỳ bà (cung huỳnh)	hát à đào	lục bát (Story of Kiều)	free	khổ giữa	f#

Table 5: Overview of sources, literary and musical metre, indicated possibility to use the bamboo slab *phách* and the reference tone as recorded (see appendix 4).

Microanalysis: 'Sa Mạc Tỳ Bà'[10]

To look at the process of transformation in detail, two examples of part 14, the first part of the final section, are examined. In the transcription, which

10 The source of this part is the so called *Long Song* (*tỳ bà hành*) which is an independent piece (*thể*, Jähnichen 1997, 1: 84) within the serious *ca trù* repertoire. It consists of twelve large sections with 8 sung lines each using a 7-7-6-8-syllable structure. The two last lines of any large section are performed without the bamboo slab. Exactly these two lines, called *sa mạc tỳ bà*, shape the 14th part of *Many Voices*. The pivotal lyrics of the *Long Song* are fixed and cannot be changed. Originally, the lyrics were written in Chinese by Đỗ Phủ in the 9th century, who created the first eight lines as the poem *Thu Hứng* (Autumn Mood), and by Bạch Cư Dị (772-846) a scholar from Thiểm Tây, later Shaanxi province of China. The Chinese text was translated into Vietnamese by Ngô Thế Vinh (1815-1869) and Phan Huy Vịnh (approx. at the same time). The quoted lyrics in *Many Voices* derive from lines number seven and eight of the third section.

was made on the spot the day after the audio recording, we first follow the order of musical lines, the rhythmic cross-points between singer and lute player and the song text.

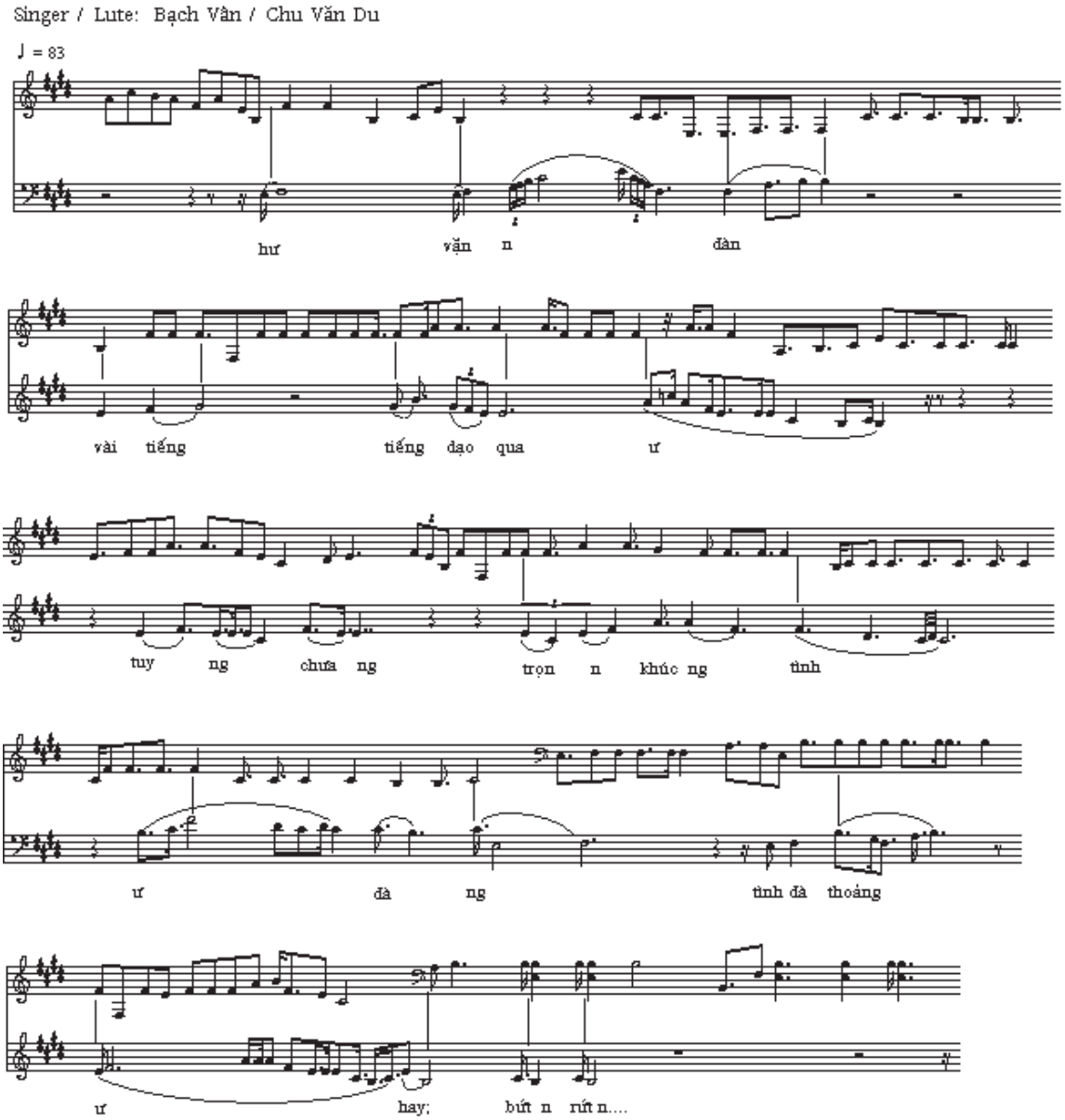


Figure 1: Lute and voice of the first example (see appendix 4) in the 14th part of *Many Voices*.

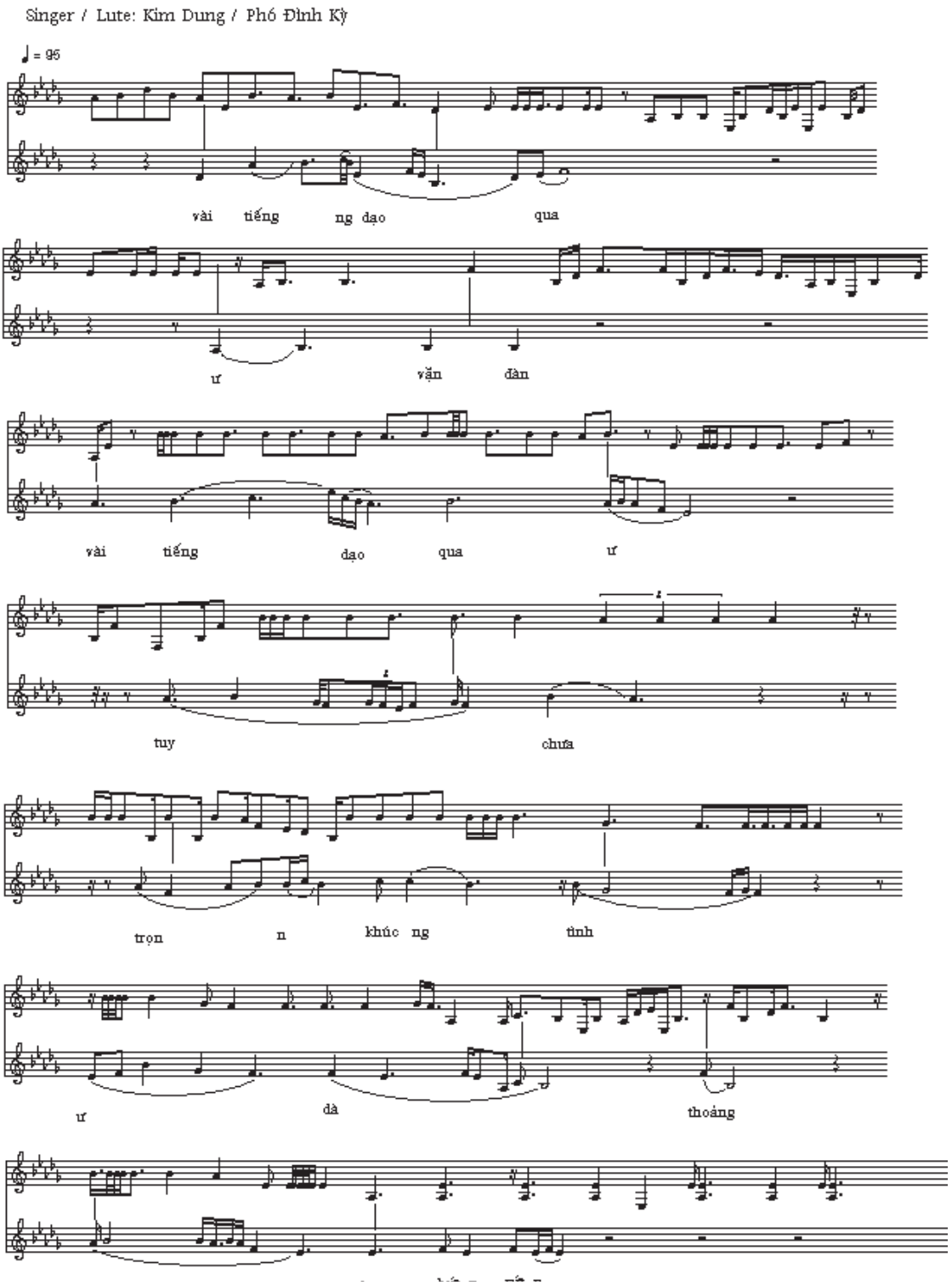
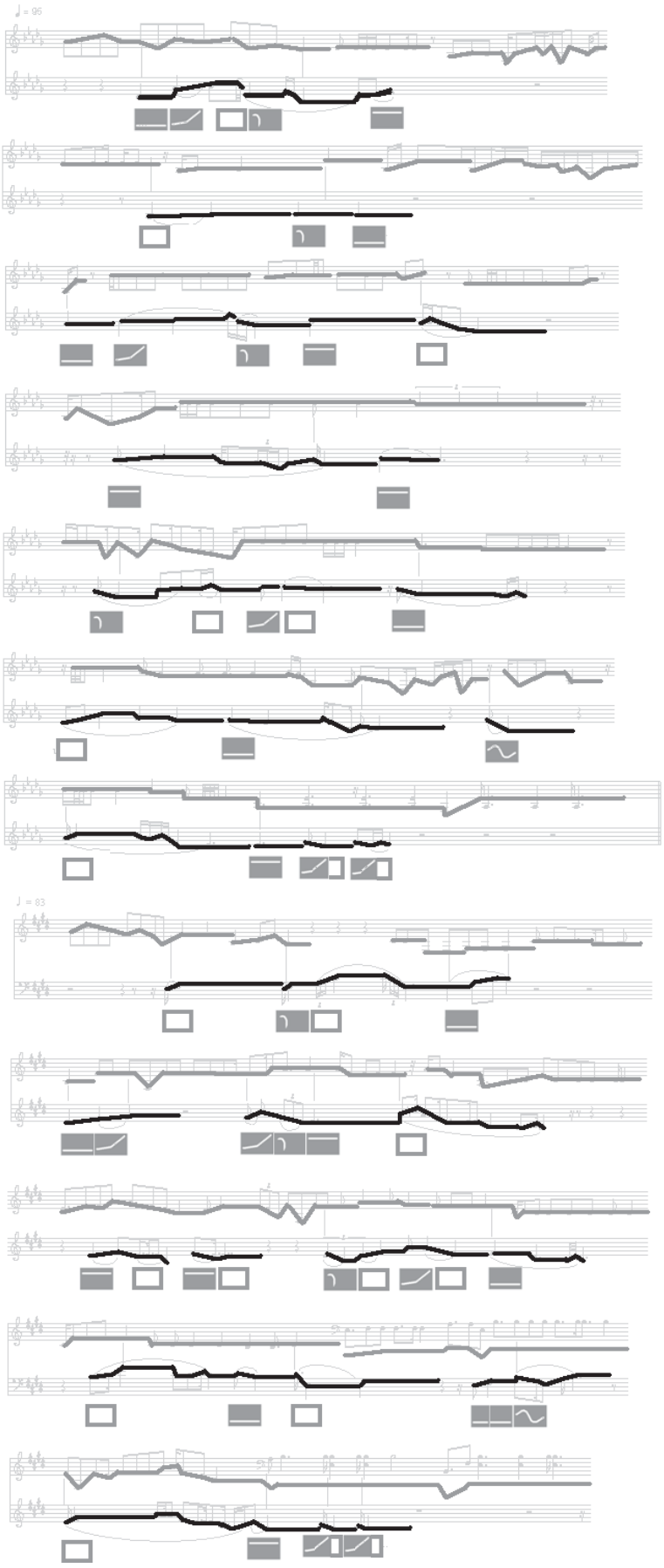


Figure 2: Lute and voice of the second example performed by Phó Đình Kỳ (1988) and Kim Dung in the 14th part of *Many Voices*.

In a second step, a close look at the contour of the singing and the lute playing in relation to the applied speech tones helps to find the right questions such as how singing lines maintain their independence as musical thoughts and with which methods they relativise speech tones to move freely within a larger melodic space.



Figures 3 and 4: Contours of lute and voice
in example 1 and 2 of *sa mạc tỳ bà*.

Both examples give a frame of individual possibilities in creating song text from pivotal lyrics. The first example goes straight into each line, while the second example starts with a so-called pre-posed repetition of the second half of the first line. The combination of speech tones remains generally similar, but varies in repetitive elements that follow the musical concept of melodic lines in dialogue with the lute.

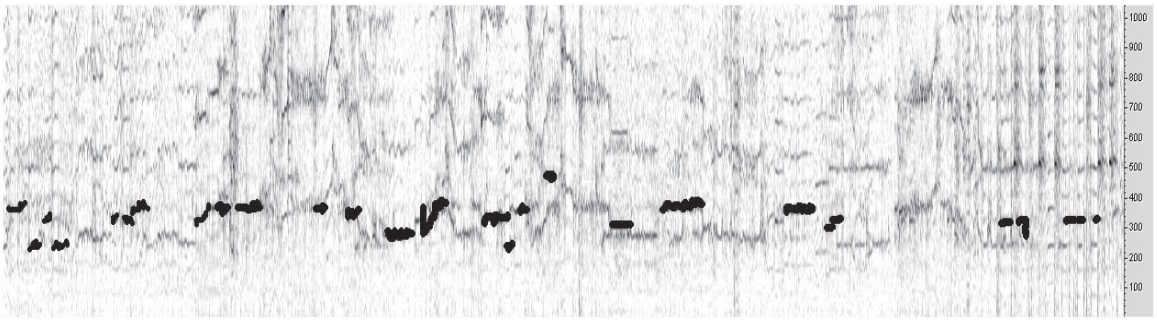


Figure 5: Spectrogram of the first example. The darker lines mark the lute part (measurement in Cents).

Though this is a clear recitation style that may imply a dominant presentation of text, contradictory melodic movements can be observed in several cases.

In the first example, second line, fifth syllable, the tonal emphasis should lead to a rising voice; however, the musical melody continues in low register.

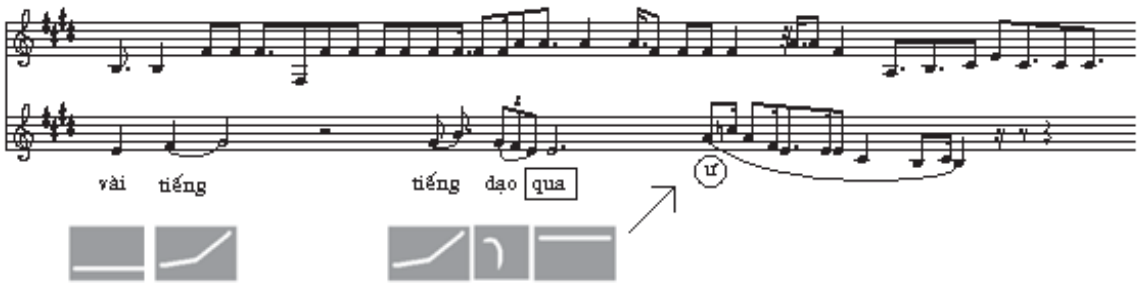


Figure 6: Transcription of the first example's second line. On the fifth syllable "qua" appears a high level speech tone that is sung on the same level as the preceding descending speech tone. A following vowel compensates the 'missing' tone level.

In the second example, the end of the second line is sung on low pitch, as is the beginning of the third line, but the third line starts a sixth above. To achieve this melodic independence, auxiliary syllables, also called "empty syllables", are used to develop a unique musical content apart from sound requirements of the text. Another method is to extend tones and thus separate tone carrying vowels and semivowels from their text meaning. The longer the tone the more possibilities are given. Finally, the lute can take over the part of indicating speech tones. Both examples apply this method.

Right in the first line of the first example, the lute clearly indicates a moving pitch of an equivalent type shortly before the singer reaches the syllable. Therefore this syllable does not need to be featured in its tonal properties. The same happens in the third line of the same example (Fig. 5).

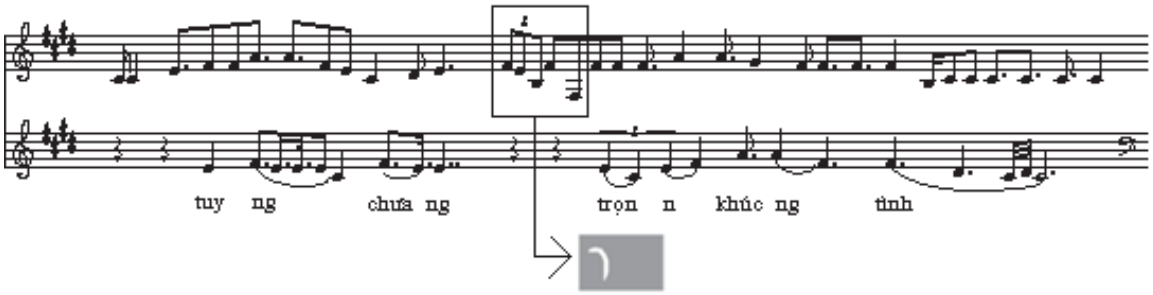


Figure 7: In the first example's third line, the lute indicates a fast moving speech tone before the respective syllable appears in the singing.

The second example shows a congruent melodic movement of the lute after the last syllable of the sixth line reflecting the speech tone, while the singer keeps on singing in low register.

Conclusion

Speech tones are important elements in vocal music as well as in the imagination of melodic aesthetics realised on musical instruments such as the lute (Jähnichen 2011) in the case of the *ca trù* repertoire. However, the often projected linearity of the relationship between speech tone and musical melody cannot be confirmed.[11] If singing is just a melodic imitation in a more or less rhythmic frame, music would not have its own expressive space. To explore this expressive space, different ways of approaching text structures are applied and creative ideas are realised to free melodies from their textual limitation.

As the case of *Many Voices* shows, vocal traditions of various sources are jointly transformed and shaped in an embracing way that does not destroy the clarity of language coming from different historical times.

11 Feld and Fox stated already in 1994 that "taken as a whole, the literature has tended toward programmatic speculation and suggestive analogies directed from linguistic structures to musical ones" (Feld & Fox 1994: 26). Similar remarks can be found in Hoàng (1989) and Edmondson (1998), but also in many popular literature put on websites by renowned scholars such as Nguyễn Vĩnh Bảo or Trần Văn Khê (cf. e.g. <http://vinhbao.theonly1.net/South3.htm>, last accessed 14 August 2012).

Please visit http://www.phonogrammarchiv.at/7DE709JKSAF for additional appendices.

References

Brunelle, Marc. 2009. "Northern and Southern Vietnamese Tone Coarticulation: A Comparative Case Study". *Journal of Southeast Asian Linguistics* 1: 49–62.

Clayton, Martin. 1996. "Free Rhythm: Ethnomusicology and the Study of Music without Metre". *Bulletin of the School of Oriental and African Studies* 59 (2): 323–332.

Đỗ, Bằng Đoàn & Trọng Huề Đỗ. 1962. *Việt Nam ca trù biên khảo* [Studying Vietnamese *ca trù*]. Saigon: Văn Khoa.

Edmondson, Jerold A. 1998. "Tones and voice quality in modern northern Vietnamese: Instrumental case studies". *Mon-Khmer Studies* 28: 1–18.

Feld, Steven & Aron A. Fox. 1994. "Music and Language". *Annual Review of Anthropology* 23: 25–53.

Hoàng, Thị Châu. 1989. *Tiếng Việt trên các miền đất nước: Phương ngữ học*. Hanoi: Khoa học xã hội.

Jähnichen, Gisa. 1995. Cuộc thử nghiệm về Hát ả đào [Essay on hát ả đào]. Re-edition in: Nguyễn Đức Mậu (ed.). 2003. *Ca trù nhìn từ nhiều phía* [Ca trù seen from various perspectives]. Hanoi: Nhà xuất bản Văn hóa-Thông tin, 346–388.

Jähnichen, Gisa. 1997. *Studien zu traditionellen vietnamesischen Instrumentalpraktiken des Hát ả đào und des Ca vọng cổ*. (Schriften und Dokumente zur Politik, Wirtschaft und Kultur Vietnams, 7). 2 vols. Berlin: Verlag der Deutsch-Vietnamesischen Gesellschaft.

Jähnichen, Gisa. 2008. "Teaching *Ca Trù* and Teaching Humility: Preservation and Its Subjective Conditions within Local Traditions". In: Emby, Zahid (ed.). *Turning Perspectives on South East Asian Music Practices*. Serdang: UPM Press, 35–57.

Jähnichen, Gisa. 2011. "Uniqueness Re-Examined: The Vietnamese Lute Đàn Đáy". *Yearbook for Traditional Music* 43: 147–179.

Lê, Thị Bạch Vân. 2008. "*Ca Trù* in Hanoi: Reality and Some Solutions". In: Vietnamese Institute for Musicology (ed.): *Proceedings of the International Scientific Conference Ca Trù Singing of the Việt People*. Hanoi: Vietnamese Institute for Musicology, 281–301.

Ngô, Linh Ngọc & Văn Phú Ngô. 1987. *Tuyển tập thơ ca trù* [Collection of *ca trù* poems]. Hanoi: Nhà xuất bản Văn hóa.

Nguyễn, Du. 1995 [1825]. *Truyện Thúy Kiều*. Saigon: Nhà Xuất Bản Văn Hóa Thông Tin.

Nguyễn, Xuân Diện. 2007. "Thể cách ca trù trong ghi nhận của nguồn tư liệu Hán Nôm" [The genre *ca trù* in written *Hán Nôm* sources]. *Tạp chí Hán Nôm* 3 (83): 41–50.

Nguyễn, Xuân Khoát. 1987. "Giới thiệu lối hát ca trù" [Introduction to *ca trù*]. In: Nguyễn, Xuân Khoát (ed.): *Trên con đường âm nhạc*. Hanoi: Nhà xuất bản Văn hóa, 52–70.

Norton, Barley. 1996. "*Ca Tru*: A Vietnamese Chamber Music Genre". *Nhac Viet: Journal of Vietnamese Music* (Special Issue) 5: 1–103.

Norton, Barley. 2009. *Songs for the Spirits: Music and Mediums in Modern Vietnam*. Urbana & Chicago: University of Illinois Press.

Phạm, Duy. 1972. *Đặc khảo về dân nhạc ở Việt-Nam*. Saigon: Hiện Đại.

Stock, Jonathan P.J. 1999. "A reassessment of the relationship between text, speech tone, melody, and aria structure in Beijing Opera". *Journal of Musicological Research* 18 (3): 183–206.

Trần, Văn Khê. 1982. *Einführung in die Musik Vietnams*. Wilhelmshaven: Heinrichshofen.

Personal communication, lessons:

Chu Văn Du: December 1992 – September 1995. Hanoi, Khâm thiên.

Đình Khắc Ban: December 1992 – March 1993. Hanoi, Ô chợ dừa.

Lê Thị Bạch Vân: October 1992 – September 2008. Hanoi, various places.

Phó Đình Kỳ: November 1988 – September 1995. Hanoi, Khâm thiên.

Studying Tone and Singing in the Laboratory

(Murray Schellenberg)

This paper provides a case study of some recent work on tone realization in singing in both Cantonese and Mandarin in a laboratory setting. The individual results have been reported in greater detail elsewhere (Schellenberg 2011, 2012a, 2012b), so only brief summaries will be included here. In this paper particular attention will be paid to some of the reasoning behind the choices made in both choosing to use an experimental methodology and the processes involved. It is also hoped that presenting both studies together will facilitate comparison between the two languages.

Popular vocal music in Mandarin and Cantonese provides a particularly interesting context in which to examine tone and singing. In the late nineteenth and early twentieth century there was an influential movement to reform music in China. This led to a new style of music which has influenced much of Chinese popular music since (Jones 2001, Mittler 1997). In the 1960s and 70s, this style of music, which had been predominantly Mandarin, split off into a separate Cantonese music (Choi 1990, Wong 2001). What is particularly interesting is that the Mandarin music and the Cantonese music deal with tone in very different ways.

In some of their recent work on Dinka, Ladd & Remijsen (in progress) have proposed a classification for the possible manifestations of lexical tone in singing. They divide the possibilities into structural manifestations – those which are prescribed by the musical melody – and phonetic manifestations – related characteristics of tone that are included during performance but not "required" by the prescribed music. The phonetic manifestations are further subdivided into compensation – characteristics of lexical tone which may be deliberately maintained or even exaggerated by the singer; and residue – manifestations of secondary characteristics of tone which are transferred over in a more reflexive or automatic way. This paper focuses on how experimental procedures may be used to examine phonetic manifestations.

Tone Systems

Although they share a common writing system and are often referred to as dialects of Chinese, Cantonese and Mandarin are usually regarded as separate languages by linguists. Although they are genetically related they are not mutually intelligible in their spoken forms. It has been claimed that they are about as distantly related as German and Swedish (Matthews & Yip 1994: 5).

Mandarin Tone System

1	High level	5-5
2	Rising	3-5
3	Fall-rise	2-1-4
4	Falling	5-1

Cantonese Tone System

1	High level	5-5
2	High rising	3-5
3	Mid level	3-3
4	Low falling	2-1
5	Low rising	2-3
6	Low level	2-2

Figure 1: The tone systems of Mandarin (left) and Cantonese (right).

They are both tone languages which have inherited their tone systems from a common ancestor, but they have evolved differently. The tone systems are outlined in Figure 1. The numbers in the left column of each chart are the standard numerical labels; the tones are often referred to simply by these numbers (e.g. tone 1, tone 4). The number sequences on the right are descriptions according to Chao's (1947) descriptive system. Each number represents a roughly equidistant pitch step with 1 as the lowest and 5 as the highest. A sequence such as 2-5 would represent a tone that starts not quite at the lowest part of the tone frequency range and then rises to the top of the range. Mandarin has one level tone and three contour tones; Cantonese has three level tones, two rising tones and one falling tone (although the falling tone appears to be undergoing a change, being frequently marked by the presence of creaky voice in modern Cantonese; see Lam & Yu 2010).

Structural Manifestations of Tone in Mandarin and Cantonese

Earlier studies have examined how and to what extent Mandarin and Cantonese express tone structurally in music. Cantonese has lexical tones well represented in vocal music whereas Mandarin does not. Chan (1987) compared both languages and noted the striking dissimilarities: using a tone-shape / note-shape comparison, she found that Cantonese reflects the matching tone patterns approximately 90% of the time but that Mandarin does so only about 35% of the time. Wong & Diehl (2002) found that Cantonese music conflates the six tones into three levels based on the final pitch target and that, when limited to these levels, the structure of the music matches the tone structure of the spoken language very strongly (95% of the time). That is, tone contours are eliminated and the music reflected the relative levels of the tones. This is in line with Chan's (1987) observations and also matches the principles of the structural representation of tones found in Cantonese Opera (Yung 1983, Zhang 2011) although Cantonese Opera operates with four tonal levels, adding a separate level for the low-falling tone.

The tonal representation in Mandarin music has received much less formal attention, but its paucity has been occasionally remarked (Chen 2007, Chao 1956, for example). Other than Chan's (1987) study the only other systematic investigation is Wee (2007) which finds that there is a tendency to match in positions where metrical prominence in music corresponds with metrical prominence in speech.

Very little has been done to look at phonetic manifestations of tone in these two languages. Chan (1987) observed that contour information was added in by the singers during performance of Cantonese songs, and Chao (1956: 57) suggested that singers in Mandarin may use grace notes (brief extra-metrical notes) "in order to 'smuggle in' the tone, if not already suggested in the main melody", but systematic study is needed to examine the question more carefully.

Why study this in the lab?

There are, of course, numerous ways in which such an analysis can be carried out, all of which have their advantages and shortcomings. The choice to adopt a designed laboratory experiment was made in this case for a number of reasons. As the focus of the investigation was the phonetic manifestation of tone, it was necessary to eliminate other potential influences on fundamental frequency (F0)[1] and duration. Three potential phonetic confounds are the intrinsic F0 of vowels, the influences of onset consonants on vowel F0 and the intrinsic duration of vowels.

The intrinsic F0 of vowels is the natural tendency in spoken language for high vowels like [i] and [u] to have a higher F0 than low vowels such as [a], a phenomenon first observed in German by E. A. Meyer in 1896–97 (Whalen & Levitt 1995 provide a comprehensive history). Intrinsic F0 differences have been found in a wide variety of languages; Whalen & Levitt (1995) survey 58 studies involving 31 languages, all of which exhibit this phenomenon. Their analysis includes ten tone languages in which intrinsic F0 is observed as a difference within individual tones; a high tone [i] or [u] will have a higher F0 than a high tone [a].

Similarly, the consonant that appears in front of a vowel can have a noticeable effect on the F0 of the vowel. Early work on this was carried out by House & Fairbanks (1953), but the topic has continued to receive considerable attention (see Hombert et al. 1979 for its role in the development of tones). As an example, vowels after voiceless stops are higher in F0 than those following voiced stops. Hombert et al. (1979) analyse the pitch contours associated with different onsets showing that a consonant can strongly influence the shape of the subsequent vowel's contour.

A third potential confound is intrinsic vowel duration. Lehiste (1970) found that low vowels like [a] are intrinsically longer than high vowels like [i] or [u].

1 Fundamental frequency is the acoustic correlate of pitch, which is an auditory phenomenon.

Like intrinsic F0, she found intrinsic vowel duration to hold across languages, including tone languages. There are also potential confounding influences from the music. The basic physiology of singing affects the F0 contour of notes so performance practices such as *portamenti* (sliding between pitches) influence the shape of the following note (Fujisaki 1981).

In order to counter these potential confounds it is necessary to look at syllables with the same segmental structure (viz. the same sounds in the same order) but different tones[2] set to the same melody. This will ensure that issues such as those discussed above will be held constant and any differences may be more confidently assumed to be due to the one factor that is different, namely, tone. It is possible, of course, to counter these confounds in other ways. Ladd & Remijsen (in progress), for example, used large numbers of naturally occurring samples with great success. The extremely high number of tokens allows the potential confounds to be "averaged out", as it were. Circumstances often dictate the choice of methodology. In the studies examined here there happened to be a fairly large number of subjects easily available who could all learn and perform the same song. The musical styles also allowed for the creation of special songs that would pass as "natural" songs.

Quantitative versus Qualitative Analysis

The studies discussed here raise the issue of quantitative analysis versus qualitative analysis and provide a good example of the differences. Cantonese music reflects lexical tone to a very high degree, being one of the highest of any language studied (Chan 1987, Wong & Diehl 2002, Schellenberg 2012a) whereas Mandarin music reflects it to a very low degree (Chan 1987). Qualitatively speaking, we would say that Cantonese music reflects lexical tone and Mandarin music does not but this does not tell the whole story. An interesting fact worth noting is that some Cantonese speakers actually use a metaphor of "rape" to describe Cantonese melodies that do not match the lexical contour.[3] Although this is a rather unpleasant metaphor, the fact that such a metaphor exists suggests that these mismatches occur in Cantonese and with sufficient frequency that they are noticeable. Earlier quantitative studies support this. Wong & Diehl (2002) found that the contours of Cantonese popular music correspond to the contours of spoken Cantonese 95% of the time; Chan (1987) found a correspondence of 90.7%. This means that, on average, 7.5% of the time the contours do not align. While that is certainly a small percentage of the time it still means that for every thirteen or fourteen notes in a melody there will be a transition that does

2 This is known as a minimal set.

3 I am indebted to an anonymous reviewer for pointing out this fact.

not match. This raises the question of how singers deal with these mismatches. Are they treated differently from transitions that match? A quantitative analysis lends itself well to an examination of these kinds of questions. To that end, the Cantonese song used here includes situations where both matches and mismatches occur – interestingly, singers made no distinctions.

Two Studies

Two separate experimental studies were carried out to examine the phonetic manifestation of tone in singing; one for Cantonese and one for Mandarin. As stated earlier, the results have been reported in greater detail elsewhere (Schellenberg 2011, 2012a, 2012b) so only brief summaries will be included here. Asking subjects to come into a lab to sing is fairly unnatural; asking them to sing random sentences (the procedure which is normally used for experimental linguistic studies) would produce an even more unnatural situation and could possibly lead to questions about the validity of the study. Trying to find a pre-existing song that includes a minimal set by tone with each word set to the same musical melody is very difficult if not impossible, but it was strongly felt that the stimuli should be presented in the context of a song. In the two studies discussed here this problem was dealt with by producing the songs especially for the studies.

In the Cantonese study the song was an adaptation of a melody written by Patrick Wong for a study he carried out with Randy Diehl (Wong & Diehl 2002). This melody was used for a perception study on tones and consisted of a musical setting (according to the rules laid out for text setting in Cantonese) of the phrase *ha yat go zi hai* (—)[4] 'the next word is (—)'. The musical phrase was adapted to provide matching contexts for different tones resulting in two slightly different phrases which were strung together in the form AABBAABBA, a concluding phrase was written by the author, and the whole thing was presented to the subjects as a children's song. The song is shown in Figure 3. The lyrics translate as

> The first word is (—),[5] the second word is (—), the third word is (—);
> the fourth word is (—), the fifth word is (—), the sixth word is (—);
> the seventh word is (—), the eighth word is (—), the ninth word is (—).
> We are all words.

Word	Tone	Description	Character	Gloss
Si	1	High; level (5-5)	師	'teacher'

4 The tones for the carrier phrase are *dai6* (number) *go3 zi6 hai6* (word). It should be noted that the first word *ha* in Wong's original phrase is tone 6; the same as the word *dai* in the sentence used in this experiment.

5 The literal translation is 'number one word is (—), number two word is (—)', and so on. Cantonese uses the cardinal numbers in this situation.

Si	2	High; rising (3-5)	史	'history'
Si	3	Mid; level (3-3)	嗜	'to try'
Si	4	Low; falling (2-1)	時	'time'
Si	5	Low; rising (2-3)	市	'market'
Si	6	Low; level (2-2)	豉	'fermented beans'

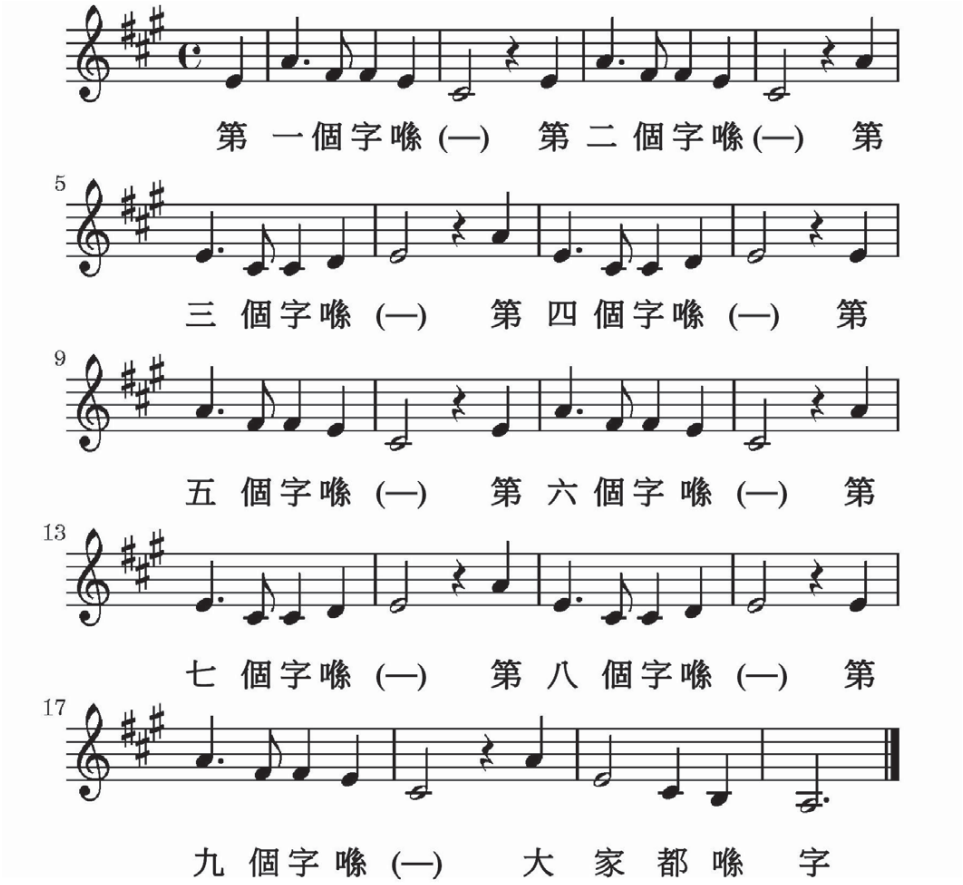


Figure 3: Score of the Cantonese song.

Subjects were told that in this song the words marked by the dash are traditionally chosen by the singers and that a further challenge is traditionally added by changing the order of the numerals: counting backwards or choosing first the even numerals and then the odd numerals; but that in the study both the word choice and the order of the numbers would be generated randomly by the computer (which was, in fact, the case). The target words were a minimal set on the syllable *si* as shown in Figure 3. A further selection of 12 distractor words was added so that two repeats of the song included all 6 target words. Randomisation of stimuli is typical of this kind of linguistic experiment as a counter to what are known as order effects. Order effects are confounds that can arise out of having the order of presentation never vary; an example of an order effect would be a fatigue effect where the stimuli are so predictable that the subject stops paying attention to what he or she is doing. Although it

proved successful the randomisation produced a somewhat unnatural situation for singing, and was abandoned in the Mandarin study.

There was no pre-existing melodic phrase available for Mandarin so the song for this study was produced from scratch. The first task was to find a minimal set that could be realistically worked into a carrier phrase[6] that had some semantic sense. This meant that all the words needed to be the same part of speech. The minimal set on the syllable *shi* proved the best choice. This set had the added advantage of starting with a voiceless consonant which meant the singer would have a brief period without glottal vibration which would provide a break from the preceding note. The words are given in Figure 4. These words were inserted in the carrier phrase "I opened my eyes and saw ____" which was then worked into a poem by two native Mandarin speakers (Chenhao Chiu and Yuan Lu) who are also trained linguists. The lyrics translate as follows:

As I was walking down the road
I couldn't see what was around me
I was so tired
I opened my eyes and saw the city

As I was walking through the city
I couldn't see what was around me
I was so alone
I opened my eyes and saw the teacher

As I was walking with him
I couldn't see what was around me
I miss her so much
I opened my eyes and saw my love history

As I was walking through the past
I started to see what was around me
I am so afraid
I closed my eyes but saw only my time

The target syllables appear as the second half of the final word of each stanza: city, teacher, love history, time. There are, of course, issues that arise with the writing of a poem. Two concessions to poetic constraints were made. The first was in the use of disyllabic words. To be believable the poem needed to

6 A carrier phrase is a sentence in which the words to be studied can all be inserted and still end up with an understandable sentence. Typical carrier phrases are "the next word is ____" or "I will say the word ____". They are usually semantically empty.

be written in modern Mandarin which has a strong preference for disyllabic words. It was not possible to come up with a minimal set that used the same first syllable, but it was possible to control the tone of the first syllable; the target syllable is the second member of a compound word where the first syllable carries tone 2 except for the tone two word (昔時xishi 'time'), which has a tone one/tone two sequence (see Figure 4). The second concession was the structure of the carrier phrase in the final line of the poem: it is not an exact repeat of the other stanzas. It was very strongly felt by the lyricists that the poem did not read well with a straight repetition of the carrier phrase, so their recommendation was followed.

Word	Tone	Description	Character	Gloss
wushi	1	High level	吾師	'teacher'
xishi	2	Rising	昔時	'time'
qingshi	3	Fall-rise	情史	'love history'
chengshi	4	Falling	城市	'city'

Figure 4: Target words for the Mandarin study

The poem was then set to music by the author in the style of Shanghai pop/jazz from the 1930s and 40s. The particular musical style was chosen for a number of reasons: it was felt by the lyricists that the style would work well with the lyrics; it is a style which is relatively familiar to most Chinese speakers but is distant enough in time that modern-day Chinese speakers would not be surprised to find a song of this type with which they were unfamiliar;[7] it is a style of music with which the author is very familiar and it is also quite formulaic musically. Chen (2007) provides a detailed analysis of this style of music and includes a very thorough description of the formula. The song was checked by several native Mandarin speakers familiar with this style of music and was judged to be acceptable. Each stanza is repeated to the same melody, and the two final lines are repeated at the end to allow for a closing cadence. The song is given in Figure 5. There is a slight discrepancy between the melody of the final line of the last stanza and the three previous stanzas. In order to fit in the coda of the two closing lines it was necessary to shorten the last note of the last stanza (the first note of measure 32).

7 A suitable comparison in Western music might be big band music from about the same era. Most people in North America are familiar with the style and know the top singers and some of the top hits, but there is a lot of material known only to aficionados.

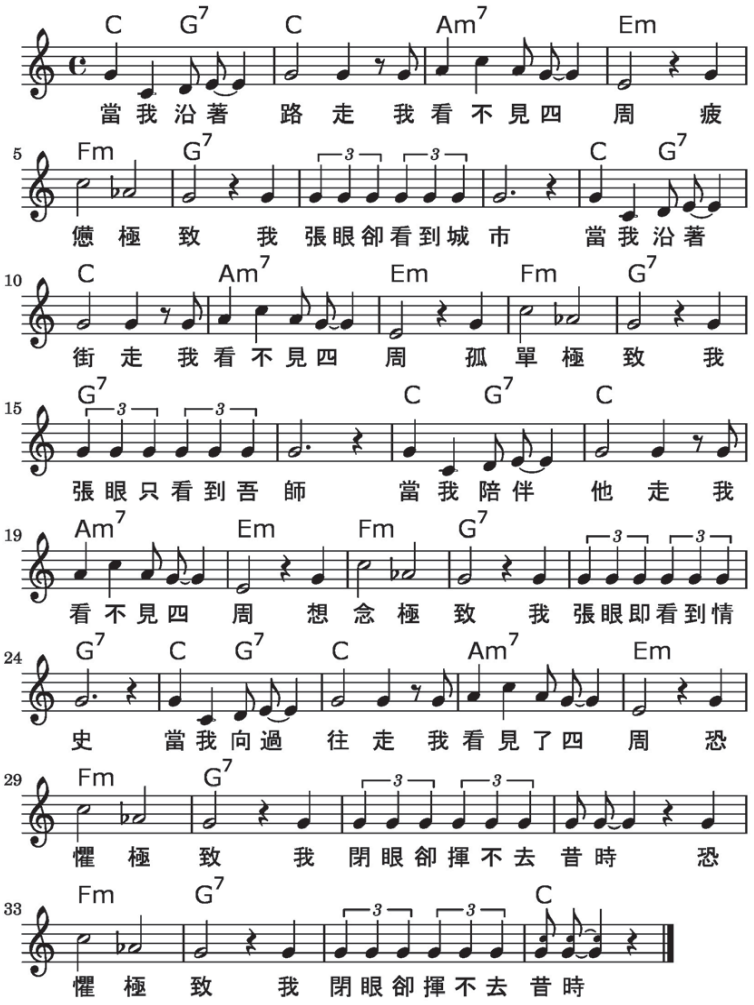


Figure 5: The Mandarin song

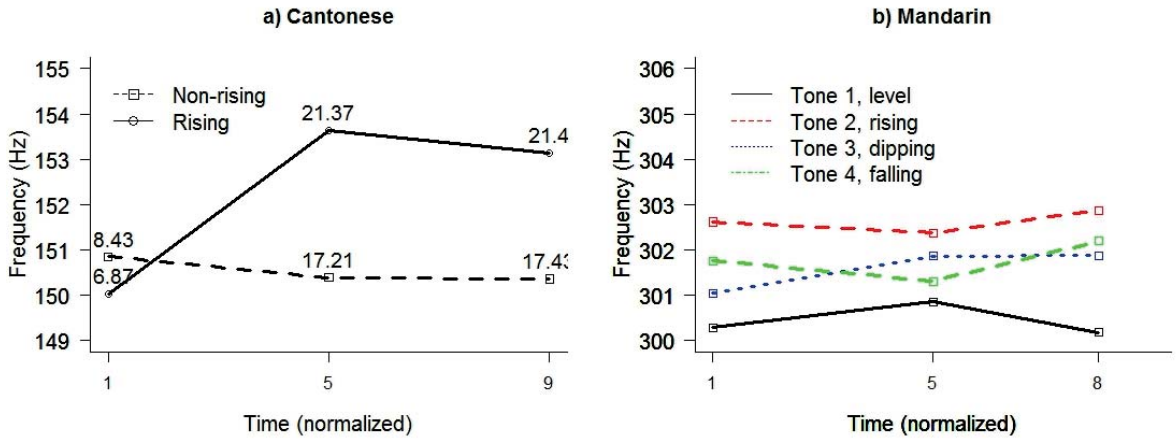
In both studies the subjects came to the lab to participate. There were twelve subjects for the Cantonese study and ten for the Mandarin. The song was presented as a karaoke song with the words presented on a computer screen using the computer program *E-prime* (Schneider et al. 2002). The subjects learned the songs by listening to a recording of the music (without words) and, in the Mandarin study, watching the words on the computer screen. Subjects were allowed to listen to an instrumental recording of the song as many times as they wished until they felt comfortable enough to sing the song. In the Mandarin study subjects were given a choice of either traditional or simplified Chinese characters for the display and were also given a choice of three different keys in which to sing. This way they could sing in a comfortable part of their range with less risk of either strain from singing too high or creak from singing too low. In the Cantonese study the singers were recorded singing a capella, but in the Mandarin study there was an accompanying track that they sang along to, produced in the computer program *Garage Band* (Apple Inc. 2012). The accompanying track was played on small speakers situated just behind the computer;

the subjects were recorded using a head-mounted directional microphone which prevented the accompanying track from being recorded. They then sang the song several times (twelve times for the Cantonese and six for the Mandarin; this allowed for six samples of each token for each language) and were recorded using the computer program *Audacity* (Mazzoni & Dannenberg 2006).

After recording, the sound file was loaded into the program *PRAAT* (Boersma & Weenink 2010) and the target words segmented out and extracted as individual .wav files. The files were examined individually and any errors corrected.

Summary of Findings

Measurements from the .wav files were transferred to the statistical program *R* (R Development Core Team 2008) for statistical analysis. Only a broad summary will be presented here; the interested reader is referred to Schellenberg (2011, 2012a and 2012b) for more detailed presentations of the statistical results. The phonetic measurements examined in these studies are slope and duration. Slope was chosen to examine the suggestion that elements of contour are added in by the singers (Chan 1987, Chao 1956). Duration was chosen because it is well known to be a robust secondary feature of tones in both Mandarin (Xu 1998) and Cantonese (Kong 1987).



1. Slope

Slope is the simple mathematical calculation of rise over run that tells how steep a line on a graph is. To calculate slope for this study, fundamental frequency measurements were taken at eleven equally spaced intervals across the duration of the vowel in each target word. The first and last points (0 and 10) were excluded to help eliminate any effects from external context. The slope was extrapolated across a number of different combinations of time points: the whole vowel (1-9), the first half (1-5), the last half (5-9) and across the midpoint of the vowel (3-7).

Extrapolation was necessary to remove any influence from vibrato. Figure 6 shows the graphs for mean slopes of the first and last halves of the note for Cantonese in (a) and Mandarin in (b). The Cantonese graph has the two rising tones combined and the four non-rising tones combined; the Mandarin graph shows the mean for each tone individually. Error bars have been omitted for ease of reading (but the standard deviations have been included in the Cantonese graph). The distinction in Cantonese is particularly striking; the rising tones show a marked rise in the first half of the note. This difference in slope is statistically significant ($\beta = 0.9158$, $t = 4.791$, $p < 0.0001$*) whereas the difference between the two for the second half is not ($\beta = 0.0349$, $t = 0.265$, $p = 0.7910$).[8]

In contrast, there is no statistically significant difference between the slopes in Mandarin (details in Schellenberg 2012b).

2. Duration

Means of the raw duration values for each tone were compared. There were no significant differences in either Cantonese or Mandarin; durational differences do not appear to be maintained in singing in either language. The one exception was tone 2 in Mandarin which is the syllable set on a shorter note (the first note in measure 32 in Figure 5) which was significantly shorter than each of the other three tones.

Conclusions

Controlling syllable structure and musical setting in these experiments made it possible to minimise various potential confounds. The use of this kind of laboratory-based experimental methodology provided strong evidence to support earlier observations regarding the presence of an added rising contour for rising tones in Cantonese. This, in turn, provides support for Ladd & Remijsen's (in progress) notion of phonetic compensation. Singers

8 Schellenberg (2011) discusses the statistics in greater detail.

in Cantonese are, indeed, adding in extra contour information regarding tone while they sing, while singers in Mandarin appear not to be.

Equally interesting are the negative results for duration. At least in Cantonese and Mandarin, singers do not appear to carry over the durational aspects of tone into singing. The reflex associated does not appear so strong as to be automatic; the singers can "turn off" aspects of tone in deference to the prescribed music.

The results raise the question of comprehension; a series of experiments are in progress to test whether these sung syllables are understandable to native listeners. The prediction is that Cantonese listeners will have more advantages when listening than Mandarin listeners and should show better results in perception tasks based on the stimuli produced by the singers in these two studies.

Experimental methodologies are certainly not the only way to examine tone in singing, but they provide an excellent way to get at small details that may otherwise be unclear or difficult to find with other methodologies.

References

Apple Inc. 2012. *Garage Band '11 (version 6.0.5)*. [computer program, Mac OS X 10.6.8.]

Boersma, P. & D. Weenink. 2010. *PRAAT: doing phonetics by computer (version 5.2.03)*. [computer program, retrieved 19/11/2010 from <http://www.praat.org/>]

Chan, M. 1987. "Tone and melody interaction in Cantonese and Mandarin songs". *UCLA Working Papers in Phonetics* 68: 132–169.

Chao, Yuen Ren. 1947. *Cantonese Primer*. Cambridge, MA: Harvard University Press.

Chao, Yuen Ren. 1956. "Tone, intonation, singsong, chanting, recitative, tonal composition and atonal composition in Chinese". In: Halle, M. et al. (eds.). *For Roman Jakobson: Essays on the occasion of his sixtieth birthday, 11 October 1956*. The Hague: Mouton, 52–59.

Chen, S.W. 2007. *The music industry and popular song in 1930s and 1940s Shanghai: a historical and stylistic analysis*. PhD thesis, Stirling University.

Choi, Po-king. 1990. "Popular culture". In: Wong, R.C. & J.Y.S. Chang (eds.). *The Other Hong Kong Report 1990*. Hong Kong: The Chinese University Press, 443–468.

Fujisaki, H. 1981. "Dynamic characteristics of voice fundamental frequency in speech and singing: acoustical analysis and physiological interpretations". *STL – Quarterly Progress and Status Report* 22 (1): 1–20.

Hombert, Jean-Marie, John J. Ohala & William G. Ewan. 1979. "Phonetic explanations for the development of tones". *Language* 55 (1): 37–58.

House, Arthur S. & Grant Fairbanks. 1953. "The influence of consonant environment upon the secondary acoustical characteristics of vowels". *The Journal of the Acoustical Society of America* 25 (1): 105–113.

Jones, Andrew F. 2001. *Yellow Music: Media Culture and Colonial Modernity in the Chinese Jazz Age*. Durham & London: Duke University Press.

Kong, Q.M. 1987. "Influence of tones upon vowel duration in Cantonese". *Language and Speech* 30 (4): 387–400.

Ladd, D.R. & B. Remijsen (in progress). *Singing in a tone language: Evidence from Dinka*.

Lam, H.W. & K.M. Yu. 2010. "The role of creaky voice quality in Cantonese tonal perception". *The Journal of the Acoustical Society of America* 127 (3): 2023.

Lehiste, I. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Matthews, Stephen & Virginia Yip. 1994. *Cantonese: a comprehensive grammar*. London: Routledge.

Mazzoni, Dominic & Roger Dannenberg. 2006. *Audacity (version 1.2.5)*. [computer program, retrieved 13/7/2006 from <http://audacity.sourceforge.net/>]

Mittler, Barbara. 1997. *Dangerous tunes: the politics of Chinese music in Hong Kong, Taiwan, and the People's Republic of China since 1949*. Wiesbaden: Harrassowitz.

R Development Core Team. 2008. *R: A language and environment for statistical computing*. [computer program, R Foundation for Statistical Computing, Vienna (Austria); <http://www.R-project.org>]

Schellenberg, Murray. 2011. "Tone contour realization in sung Cantonese". In: Lee, Wai-Sum & Eric Zee (eds.). *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong: City University, 1754–1757.

Schellenberg, Murray. 2012a. "The neutralization of tone-related duration differences in sung Cantonese". In: Ma, Quiwu, Hongwei Dong & Daniel Hirst (eds.). *Proceedings of the 6th International Symposium on Speech Prosody*. Vol. 2. Shanghai: Tongji University Press, 514–515.

Schellenberg, Murray. 2012b. "Tone realization in sung Mandarin". *Proceedings of the 3rd International Symposium on Tonal Aspects of Language*. Nanjing. Online: <http://www.tal2012.org/abstracts/28.pdf> (27/11/2013).

Schneider, W., A. Eschman & A. Zuccolotto. 2002. *E-Prime: User's Guide (version 1.0)*. s.l.: Psychology Software Tools. [computer program]

Wee, L.H. 2007. "Unraveling the relation between Mandarin tones and musical melody". *Journal of Chinese Linguistics* 35 (1): 128–143.

Whalen, D.H. & Andrea Levitt. 1995. "The universality of intrinsic F0 of vowels". *Journal of Phonetics* 23: 349–366.

Wong, Kee Chee. 2001. *The age of Shanghainese Pops: 1930–1970*. Hong Kong: Joint Publishing (H.K.).

Wong, P.M. & R.L. Diehl. 2002. "How can the lyrics of a song in a tone language be understood?" *Psychology of Music* 30 (2): 202–209.

Xu, Y. 1998. "Consistency of tone-syllable alignment across different syllable structures and speaking rates". *Phonetica* 55: 179–203.

Yung, Bell. 1983. "Creative process in Cantonese opera I: The role of linguistic tones." *Ethnomusicology* 27: 29-47.

Zhang, Ling. 2011. "Cantonese lexical tones in speaking and singing". In: Lee, Wai-Sum & Eric Zee (eds.). *Proceedings of the Psycholinguistic Representation of Tone Conference*. Hong Kong: City University of Hong Kong, 32–35.

From Words to Music: The Pitch and Rhythmic Structure of Japanese Vocal Music

(Rinko Fujita)

Introduction: phonetic features of Japanese

In contrast to stress accent languages such as English or German, Japanese as a pitch accent language has a stable temporal structure which is based on the mora as a time unit. The mora as a phoneme typically represents a consonant sound followed by a vowel sound, or just a vowel alone.[1] In Japanese speech, the time value of this unit is pronounced with equal length. A mora is different from the concept of a syllable in Western languages. For example, the English word "Austria" consists of three syllables [aus], [tri] and [a], but a Japanese speaker subdivides the word into six units (six moras) [o], [o], [su], [to], [ri], and [a], which corresponds to six graphemes of Japanese. Furthermore, long vowels, syllabic nasals, and geminate consonants are also moraic (Shibatani 1990: 58; Takeuchi 1999: 43).

Because of this simple phonetic structure, Japanese has a relatively small number of phonetic units (syllables) in comparison to other languages. According to the Japanese linguist Kindaichi, the total number of syllables (phonetic units) in the Japanese language is only 112, while English has over 30,000 (Kindaichi 1978: 104, 1989: 291). In addition, each mora contains individual tonal information of high or low tone which determines the lexical meaning of the word. For example, the word *hashi* (はし), which consists of two moras: if the pitch of the first mora is higher than the second one ([ha↑] [shi]), it means 'chopstick' (はし). However, if the pitch of the second mora is higher than the first one ([ha] [shi↑] はし), it means 'bridge'. If there is no pitch difference ([ha] [shi] はし), it means 'edge'. Thus, in order to deliver the textual meaning precisely, it is necessary to keep the original pitch and equivalent time unit of the words. In other words, if the time unit of the word is shortened or lengthened extremely and the pitch accent is modified, the word could change (or lose) its original meaning.

The mora as a time unit also plays an important role in the composition of Japanese poetry. It is well known that a *haiku* (俳句) consists of three lines of five, seven and five moras. This 5-7-5 structure is the foundation of Japanese vernacular rhythms; not only in poems but also in music.

Such phonetic features of Japanese are very significant in music, because they can determine how far the pitch and time units (syllables) of

1 A *mora* corresponds to Japanese syllabic script *hiragana* and *katakana*.

the original words can be modified in order to conform to the melodic and rhythmic styles of the music.

This paper focuses on the pitch and rhythmic structure of Japanese vocal music. Three different vocal delivery styles — *rôdoku* 朗読 (declamation), *ginshô* 吟誦 (reciting) and *kashô* 歌唱 (singing) — applied to the same text will be compared and the acoustic features, rhythm and melodic contours of each performance will be examined via the analysis of pitch contours and the audio signal analysis.

Previous studies on the relationship between the pitch and rhythmic structure of Japanese vocal music

In contrast to Western music, which — at least until the 20th century — was based on the same musical theories and fundamental systems, each genre of traditional Japanese music has developed its own body of theories and systems, and there are only a few single theoretical systems that transcend genre boundaries. For that reason, most existing research regarding traditional Japanese music is focused on individual genres.

Fumio Koizumi (1927–1983), however, was one of the few pioneers who attempted to reveal the musical principles underlying Japanese music. He investigated a great number of types of Japanese music and at the same time examined children's songs and folksongs as well as speech in order to find "the common roots of language itself from which all kinds of Japanese music have branched" (Koizumi 1980: 108). Regarding the relationship between the musical beats and moras in a text, he points out that the "two moras-one beat" relationship is commonly considered as a primary structure in Japanese vocal music, and this two moras-one beat structure can be subdivided into three categories (Koizumi 1984: 227f., 1997: 163f.):

1. Two moras are equivalent to one beat.
2. The first mora is shorter than the second.
3. The first mora is longer than the second.

Furthermore, he suggests that the distribution of moras on a beat can express the basic character of the respective music: The one mora-one beat structure as the basic temporal principle of speech can often be found in children's charm-like songs and religious music,[2] and also in the somewhat formulaic pattern of *heike*[3] in which the verbal information is of primary importance

2 For example *shômyô* 声明.

3 *heike*: the musical recitation of *Heike monogatari* ('The Tale of the Heike'). The term *heikyoku* and *heikebiwa* are also used for this genre. Since the Japanese medieval period, this genre has simply been called *heike* among specialists, and since the term is now widely used in Japanese literature, I will use it in this article.

in the music. The pitch contours in these types of songs (or rather musical recitation) correspond to the pitch of the original words (Koizumi 1997: 150). The second type of two moras-one beat structure — the first mora is shorter than the second — is frequently found in folk music with a religious character and working songs associated with heavy kinetic movement (activity), as well as in traditional songs like *shigin* 詩吟, which express the brave and manly character of their textual content. The third category of two moras-one beat structure — dotted rhythm or skipping movements of rhythm — conveys essentially bright and cheerful expression, which is also found mainly in children's songs.

Although Japanese phonology is based on equivalent temporal units, this does not mean that all Japanese vocal music is based only on a metric system. In fact, free rhythm is very common in all kinds of traditional genres, in which a highly developed performance practice is revealed. Japanese vocal music is traditionally categorized into two major vocal styles, namely *katarimono* 語り物 and *utaimono* 歌い物. *Katarimono* refers to musical narrative in which the textual property is emphasized, while *utaimono* refers to lyric song with an emphasis on melodic properties. Kenji Hirano (1929–1992) examined the relationship between language and music in *katarimono*, and he postulated three varieties of vocal delivery style, namely *ginshô* 吟誦, *eishô* 詠唱 and *rôshô* 朗誦 (Hirano 1990). Compared to the other two styles of musical narrative, *ginshô* is closest to speech, which is characterized primarily by the "one mora–one beat" relationship, i.e. one mora is distributed on a beat. However, in this style, in contrast to normal speech and declamation, a certain musical intonation comes into play, by which the specific intonation becomes stylized and the time values of the moras of each word are emphasized through so-called *hikinobashi* 引き伸ばし (the prolongation of moras).[4] In view of the fact that *ginshô* still does not attain a fixed pitch and its time value is prolonged in an instable way, this delivery style is less recognizable as music. Even though the vocal style of *ginshô* is closest to speech, the pitch contours in this style do not correspond to the pitch accents of the original words. According to Hirano, the vocal delivery of *eishô* and *rôshô* are considered to be a development of *ginshô*: the former is characterized by long extended melismatic notes with fixed pitch, whereas the latter consists of a relatively simple musical structure in which the pitch and time values of moras correspond almost to the original words (Hirano 1990: 35–37). Hirano remarked that the linguistic features of words have a great influence on the musical structure in *katarimono,* or rather the phonetic expression of words can be regarded as the vocal musical expression of *katarimono* (ibid.: 33).

4 Sakai Yasuko (2004) examined the rhythmic structure of prolongation in Japanese vocal music.

Similarly to most countries in the world since the 1960s, Anglo-American popular music has also influenced Japanese popular music. Yoshiaki Sato investigated how the Japanese vocal rhythm was transformed into the eight-beat rhythm of rock music (Sato 1999) and suggests that there has been an adaptation process of rock music in Japan during the last 40 years: in contrast to early popular music, in which the vocal rhythms were characterized by two moras on one beat, a four moras-one beat structure became trendy in Japanese popular music compositions in the 1970s. In order to fit in with the syncopated beat of rock music, the vocal rhythm of the text was modified; the consonants, geminate consonants and long vowels were reduced or expanded extremely. Furthermore, the original accents of Japanese were changed; the equivalent time units of phonetic character were intentionally distorted so as to imitate English articulation (pronunciation). Since the 1980s, song texts have been frequently written mixed with English words. Although they sometimes make no sense with regard to the content, only the sounds of English are important, because they appeal to young people. Under the influence of hip hop of the 1990s, the vocal rhythms returned to the four moras-one beat structure again (ibid.: 139, 185).

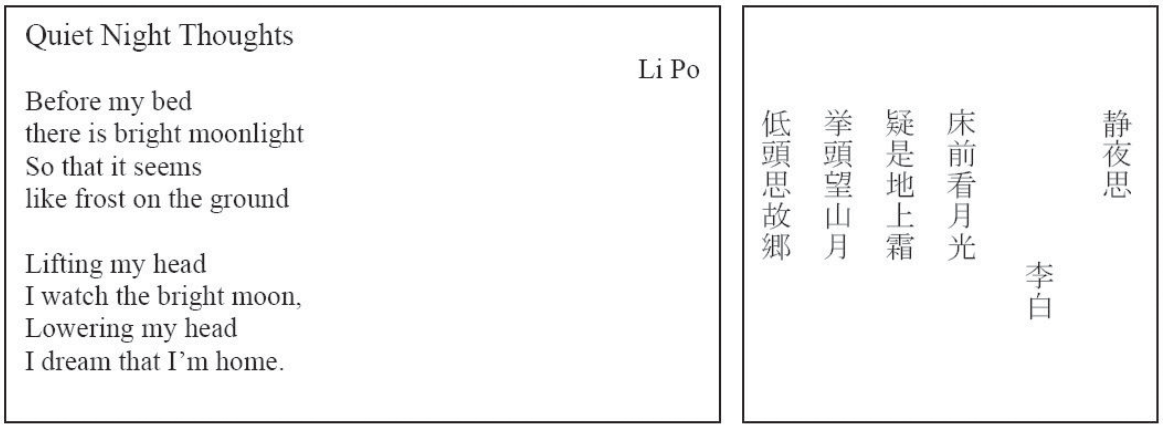
Analysis of the pitch and rhythmic structure of three different vocal delivery styles

As previously mentioned, the phonetic nature of the Japanese language is based on the equivalent time value and the low/high pitch accents. What happens to such a monotonous rhythm with the simple tonal accents of Japanese when it is sung as "music"?

In order to examine the acoustic properties, rhythmic structures, and the relationship between pitch and text two different analyses — audio signal analysis and the analysis of pitch contours — were conducted. The three recorded samples of different vocal delivery styles — declamation (cf. Terada 2003), reciting (cf. Yamamoto 1999) and singing (cf. Yanagi 1975) — within the same text were used for these purposes, during which each recorded sample was performed by a different performer.

The textual material: Chinese poetry

The textual material used for this analysis is *Quiet Night Thoughts* 静夜思 from the Chinese Tang Dynasty (618–907), composed by the famous Chinese poet Li Po 李白 (701–762), also known as Li Bai (Rihaku in Japanese); the translation is based on Cooper (1973).



The Japanese call classical Chinese poetry *kanshi* 漢詩. Ever since ancient times, Chinese poetry has been one of the popular genres of Japanese poetry, along with the *waka* and the *haiku*. Originally, *kanshi* meant classical Chinese poetry from ancient times; however, current use of the term *kanshi* also comprises poems which are composed in Japan according to Chinese prosody. *Kanshi* is written only with Chinese characters and recited with Japanese pronunciation, so that all three of the sound materials in this analysis — declamation, reciting and singing — are performed in Japanese.

Here I should clarify something about Japanese writing systems and their phonetic features: the Japanese use three basic character sets in their writing system, namely the Chinese logogram *kanji*漢字[5] and another two sets of syllabic script, *hiragana* ひらがな and *katakana* カタカナ, which represent different sounds of Japanese. Because of the different linguistic structures of Chinese and Japanese, Chinese characters were not well-suited to write Japanese; consequently, Chinese characters are used to write words of Chinese origin or words which have the same or similar meaning. And these are pronounced either with native Japanese "*kun*-reading" (explanatory reading) or pronunciation derived from Chinese "*on*-reading" (Sino-Japanese reading). Thus, a logogram can be pronounced with several different sounds. This multiple reading system is a distinctive feature of the Japanese language and makes it possible to recite Chinese poems with a Japanese pronunciation although they are written only with Chinese characters.

5 Chinese characters were brought to Japan in the 4th century AD via Korea. In the next two centuries, when Chinese books on philosophy and Buddhism made their way to Japan, the knowledge of Chinese characters spread among the Japanese aristocracy.

Reciting of Chinese poems: *shigin* 詩吟

The audio material which is used for the analysis of the reciting style is the musical genre *shigin*. *Shigin* belongs to the category of *katarimono* (musical narrative) and refers to the musical recitation of Chinese poems in Japanese. At present, *shigin* is very popular among elderly Japanese. There are nationwide networks of *shigin* amateurs.

The style of current *shigin* performance is based on the *Satsuma biwa* tradition 薩摩琵琶 (a style of musical recitation accompanied by a four-stringed pear-shaped lute). The narrative style of *Satsuma biwa* can be traced back to the 16th century in the feudal domain of *Satsuma* in the southern part of *Kyûshû*, where the *biwa* was used for pedagogical purposes to enhance the population's loyalty and fidelity to the ruler in samurai society (De Ferranti 1991: 102). As most forms of traditional Japanese music are structured as sequences of formulaic melodic and rhythmic patterns, melodies of *Satsuma biwa* also consist of sequences of formulaic melodic materials. According to the Japanese musicologist Komoda, one of the formulaic patterns called *gin'ei* 吟詠, which is used for the recitation of Chinese poems in *Satsuma biwa* performance, developed later into an independent vocal genre known as *shigin*. Because of this historical background, *shigin* used to be recited mainly among *biwa* performers until the mid-20th century (Komoda 2008: 421). The musical scale of *shigin* is generally based on a pentatonic scale called *miyakobushi onkai* 都節音階, which consists of two tetra chords combining a minor second and a major third. The reciting features are based on the prolongation of the notes and the stereotyped melodic patterns. Conventionally, the final note of each line ends on the key note of the scale.

Audio signal analysis

The primary purpose of this analysis is to examine the acoustic features of three vocal delivery styles. The audio signal analysis serves as a more precise tool for analyzing and recognizing vocal delivery, and the sound spectrogram enables a visual representation of performance.

Method: Each audio sample is divided into four segments according to the formal structure of Chinese poetry. And then the acoustic features of each segment are analyzed by using audio signal analysis software (EmapSon Version 1.01.14).

Result: Figure1-a shows the sound spectrogram of the first line of the poem's declamation. During the declamation the text was recited fluently; therefore the duration from the onset of a mora (syllable) to the onset of the following mora is obscure. In general the sound contains strong noise components; they are represented by grey shadow in the spectrogram and

can be identified particularly by the fricative sounds of the words [*sho*] and [*zé*].

In the reciting style the voice slides up into the pitch at the beginning of the performance. The vowels and syllabic nasal /N/ of the words are frequently very prolonged and ornamented with peculiar vibratos in which the pitch fluctuates irregularly with a frequency of 1 to 10Hz. Characteristic melodic ornamentations also appear in these long extended tones. The audio signals contain fewer noise components than in the declamation. [Figure1-b]

The vocalization of the singing in the example conforms to the western classical singing style. The constant vibratos on each note (amplitude fluctuation in the range of 3–5 Hz and oscillation at around 6 cycles per second) and the harmonic partials within the frequency range of a maximum of 10 kHz are characteristic for this singing style. The sound contains relatively weak noise components compared to the declamation and reciting styles. [Figure1-c]

From the rhythmic point of view, the original temporal structure of the words, which is characterized by a fairly regular pulse, is still maintained in the declamation and reciting styles, although the latter conveys an emotional impression to the listeners. In both styles, two moras are distributed equally on one beat. On the other hand, the temporal structure of the words in the singing style is obviously modified: especially the distribution of moras in the second line is modified by the so-called dotted rhythm of jumping or skipping movements. In other words, in the singing style the time value of each mora is varied and the "two moras-one beat" structure is modified. [Figure 2]

The analysis of pitch contours

As previously mentioned, in Japanese speech each mora contains individual tonal information of high or low tone which determines the lexical meaning of the word. In this section the pitch contours of each performance are examined in order to observe the distinctions among each delivery style.

Method: In order to compare the relationships between pitch and text (i.e. pitch contour) the duration and the pitch of each word are measured using digital audio editor software (Sony Sound Forge 7.0). The results are illustrated in the table, in which each cell of the table corresponds to a time-unit mora and the vertical line refers to the pitch.

Result: Table 1 shows the result of the relationships between pitch and text of three performances.[6] There are some modifications of the original

6 The text in Chinese characters is arranged in the Japanese reading system in the table.

text (marked with bold letters in the table): in the reciting style some words are added to the text (*shimo naran ka to* instead of *shimo ka to*) and in the singing style some Chinese characters are pronounced with explanatory reading.[7]

The pitch accent pattern (high-low or low-high accentuation of the words; marked in grey in the table) is very similar in both the declamation and reciting styles, i.e. the original pitch accent of the words is still maintained during the performance. In the declamation each mora was pronounced with almost the same length. However, in the reciting style some vowel and syllabic nasals /N/ which are extremely prolonged and ornamented with vibratos[8] last more than 6.5 seconds. Because of the prolonged notes which are not based on any metric system, the whole performance does not give a musical impression.

Since the third line of the four-line composition is referred to as the climax of the poem, the first half of the third line is recited in the highest pitch in the performance of the reciting style, in which the final note of each line ends on the key note of the scale. Furthermore, in the reciting style two different kinds of ornamentation are frequently used at the end of the prolonged notes. [Figure 3]

In the singing style, hardly any words keep their original pitch accent. The pitch accent of the words is forced to conform to the melodic line; consequently almost all the words lose their original accentual structure, and the time unit (mora) of the words is multiplied or divided according to the rhythm of a given music. In other words, the pitch accent is modified towards purely musical expression.

Conclusion and final remarks

The pitch and rhythmic structure of Japanese vocal music have to be observed more systematically and precisely. However, the results of this analysis have shown that the original pitch and rhythmic structure in the declamation of Chinese poems are still retained during recitation. On the other hand, the vocal style of the recitation is musically enriched by the melismatic ornamentation and vibrato on the prolonged notes. In singing, the original words of the text are forced to conform to the melodic line, and the pitch and rhythmic structure of each word is distorted. As a consequence, the individual words are less intelligible.

7 In singing, 月光 ('moonlight') is pronounced *tsuki no hikari* instead of *ge'kko*. *Tsuki no hikari* is explanatory reading of the characters月光; they share the same meaning.

8 The prolongation is marked with a capital L in the table.

In declamation and recitation, the emotional value of the text is primarily expressed through the performance. Or more specifically, in recitation the emotional value is expressed through the intensive vibratos and melismatic ornamentations on extended notes, while in declamation it is expressed through the intensity (loudness) of the voice. On the other hand, in singing the emotional value of the text is expressed through the compositional technique, in which dynamics and agogics are used as compositional tools.

Considering the two (seemingly contradictory) accounts of the phenomenon that Japanese phonology is based on equivalent temporal units and that free rhythm (non-metric structure) is very common in all kinds of traditional musical genres, how shall we comprehend the relationship between words and music? Is musical narrative, which has a tendency towards a melismatic and non-metric nature, a rudimentary stage in a lineal development from words to music?

It is a historical fact that after the medieval period in Japan, various genres of musical narratives developed alongside performing arts and a variety of vocal delivery styles began to be used. Moreover, these vocal delivery styles are very often used alternately within a composition. For example, in the puppet theatre *gidayû-bushi* 義太夫節, the dialogue and narrative as well as the songs are performed alternately. Once the spoken dialogue shifts fluently into a narrative or song, vocal delivery is no longer limited by phonetic constraints. The time units of the words can then be lengthened or shortened in line with the needs of the artistic expression of the music.

As mentioned above, Japanese vocal music is traditionally categorized into two major vocal styles, *katarimono* (musical narrative) and *utaimono* (song). The conceptual categorization implies there is no hierarchical differentiation between the two categories. Considering that a variety of highly-developed vocal delivery styles are used in musical narrative, free rhythm should be comprehended as one of the important musical devices for artistic expression, in which the words are "freed" from phonetic bounds. In Japanese musical culture it is therefore not necessary to interpret the shift from the words (speech/narrative) to song as a lineal development of music. Each musical genre of traditional Japanese music has developed its own styles of expressing aesthetic ideas through music. [Figure 4]

Figure 1: Sound spectrogram of the first Line of *Quiet Night Thoughts.*

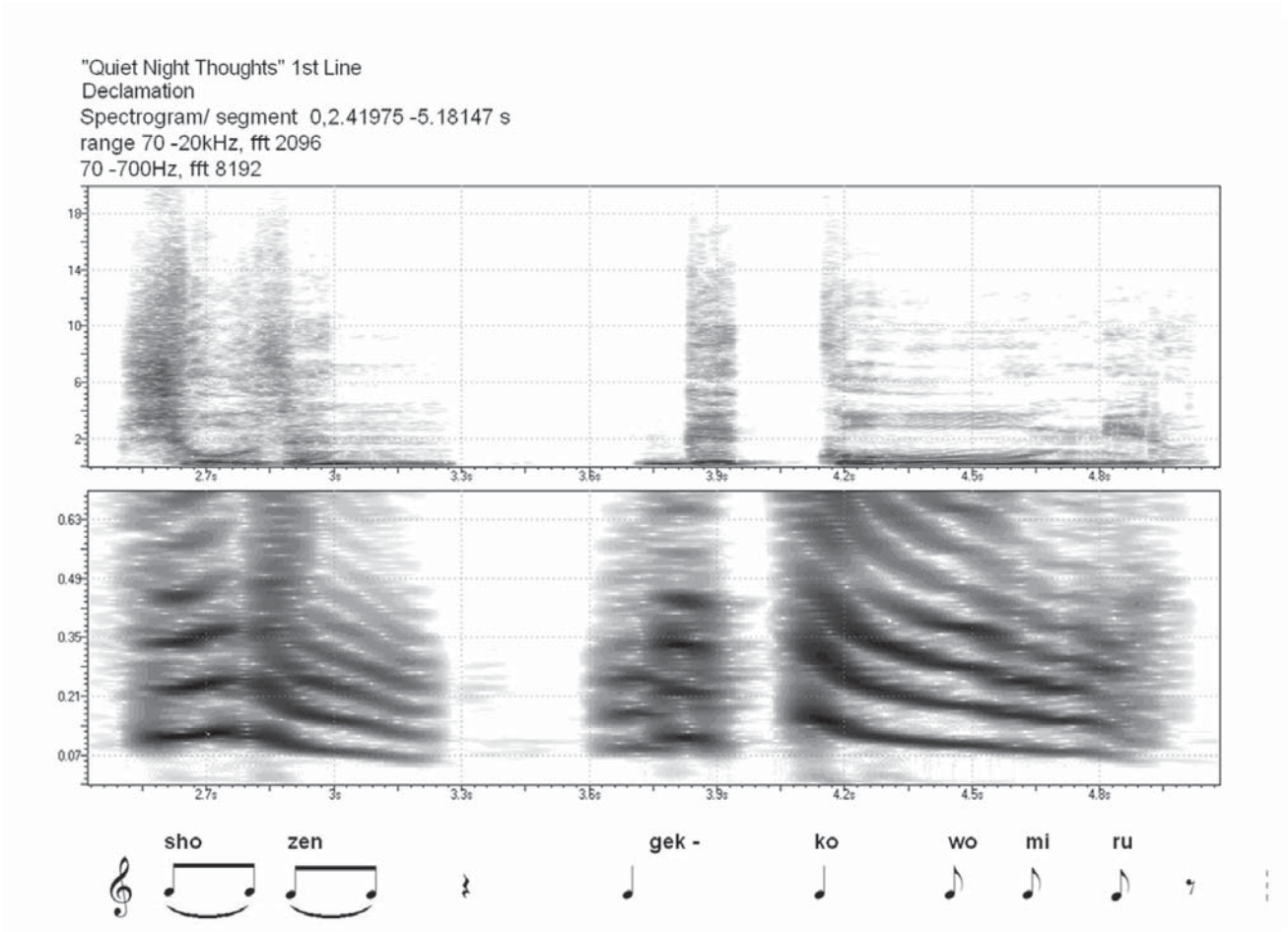


Figure 1a: Declamation

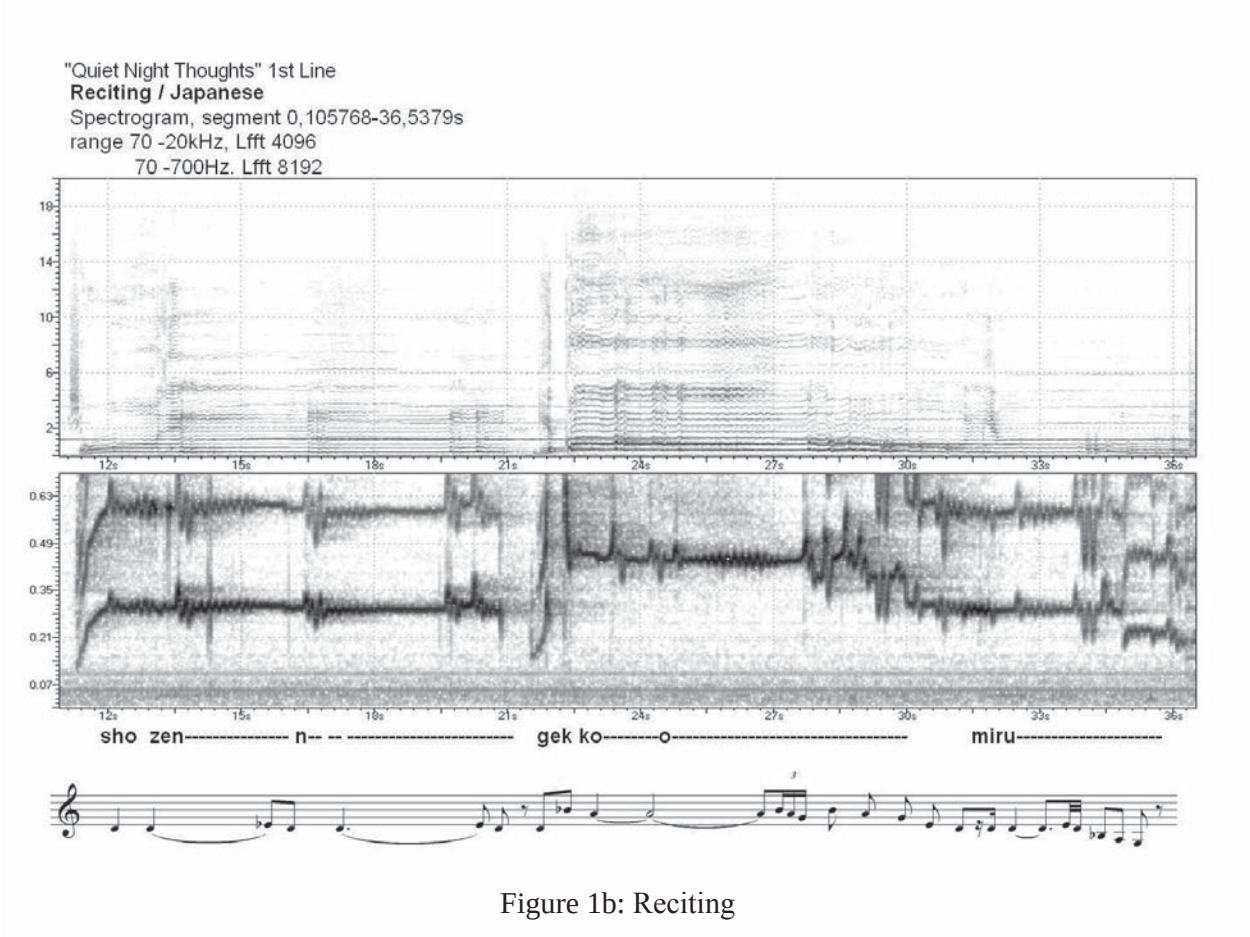


Figure 1b: Reciting

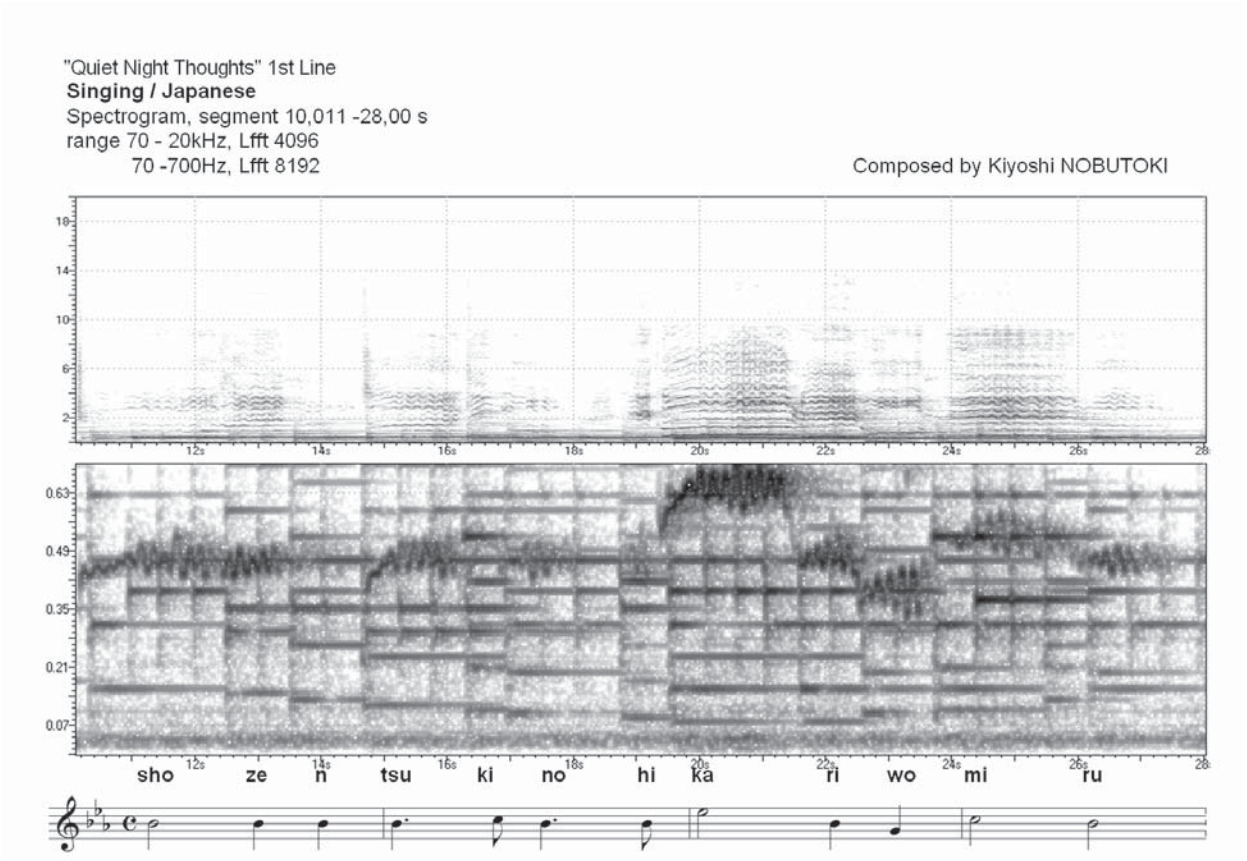


Figure 1c: Singing

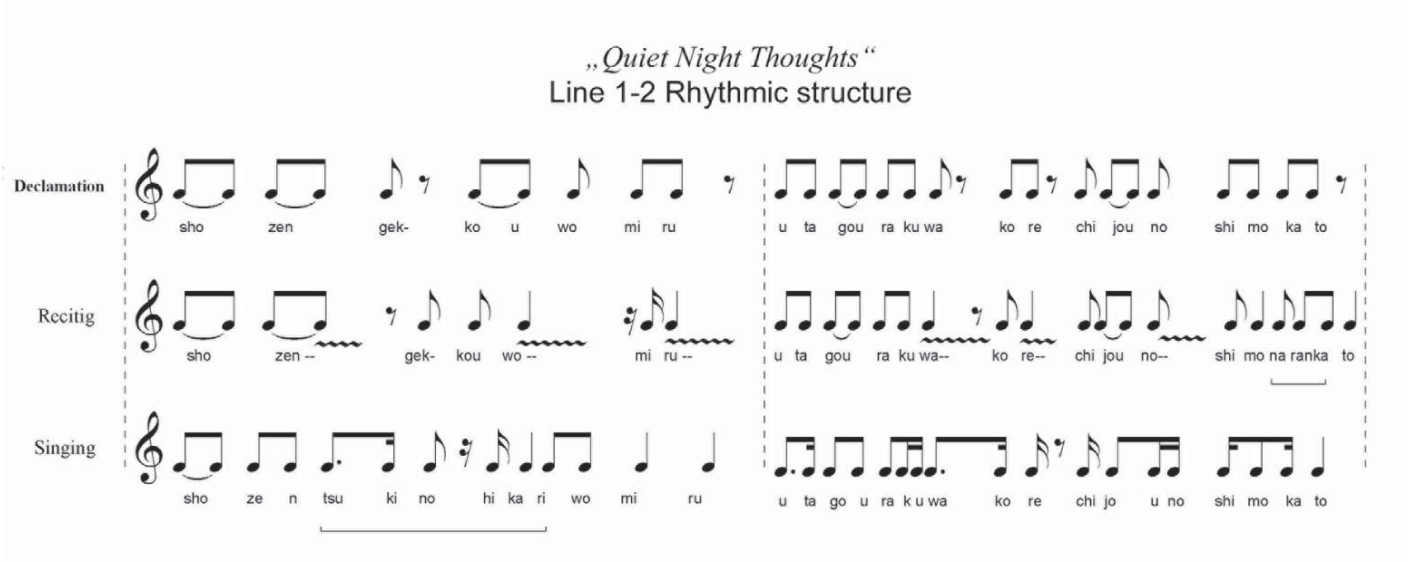


Figure 2: Mora distribution

Declamation

朗読　静夜思	Line 1									Line 2																Line 3															Line 4														
Pitch H																																																							
Pitch L																																																							
Chin. chracter	床		前		月		光		看	疑							是	地	上		霜					頭			挙		山		月		望					頭			低		故		郷		思						
phonetic unit mora	sho	u	ze	n	gek	ko	u	wo	mi	ru	u	ta	go	u	ra	ku	wa	ko	re	chi	jo	u	no	shi	mo	ka	to	ko	u	be	wo	a	ge	te	sa	n	ge	tu	wo	no	zo	mi	ko	u	be	wo	ta	re	te	ko	kyo	u	wo	o	m	o	u
number of mora	2		2		3		1		2	7							2		4		4					4			3				4					4																
time value in sec.	0,30		0,47		1,12				0,30	1,06							0,47		0,77		0,77					0,66			0,40				1,13					0,56																

Reciting

吟誦　静夜思	Line 1									Line 2																Line 3															Line 4														
Pitch B♭																																																						
Pitch A								L																																														
Pitch G																																																						
Pitch F																																																						
Pitch E♭																																																					
Pitch D			L									L															L		L					L								L										
Pitch C																																																					
Pitch B♭																																																					
Pitch A																																																						
Pitch G																																																						
Chin. chracter	床		前		月		光		看	疑							是	地	上			霜				頭			挙		山		月		望					頭			低		故		郷		思						
phonetic unit mora	sho	u	ze	n	gek	ko	u	wo	mi	ru	u	ta	go	u	ra	ku	wa	ko	re	chi	jo	u	no	shi	mo	na	ra	n	ka	to	ko	u	be	wo	a	ge	te	wa	sa	n	ge	tu	wo	no	zo	mi	ko	u	be	wo	ta	re	te	wa	ko	kyo	u	wo	o	m	o	u
number of mora	2		2		3		1		2	7							2		4			4				4			4				4					4																
time value in sec.	1,31		7,92		9,87				1,84	5,82							6,78		6,38			5,73				7,32			7,80				9,89					4,64																

Singing

歌唱	Line 1										Line 2																	Line 3																Line 4															
Pitch G																																																									
Pitch F																																																									
Pitch E♭																																																										
Pitch D																																																									
Pitch C																																																									
Pitch B♭																																																									
Pitch A♭																																																									
Pitch G																																																									
Pitch F																																																									
Pitch E♭																																																									
Pitch D																																																									
Pitch C																																																									
Pitch B♭																																																									
Chin. chracter	床	前		月			光		看	疑							是	地	上			霜				頭			挙		山		月		望					頭			低		故		郷		思						
phonetic unit mora	sho	u	ze	n	tu	ki	no	hi	ka	ri	wo	mi	ru	u	ta	go	u	ra	ku	wa	ko	re	chi	jo	u	no	shi	mo	ka	to	ko	u	be	wo	a	ge	te	sa	n	ge	tu	wo	no	zo	mi	ko	u	be	wo	ta	re	te	ko	kyo	u	wo	o	m	o	u
number of mora	2	2		3			4		2	7							2		4			4				4			3		4				4					4			3		4				3						
time value in sec.	2,32	2,28		4,05			4,97		4,16	8,16							1,88		3,03			4,13				4,14			4,62		4,44				3,72					4,69			3,69		5,17				6,45						

Table 1: Relationship between pitch and text.

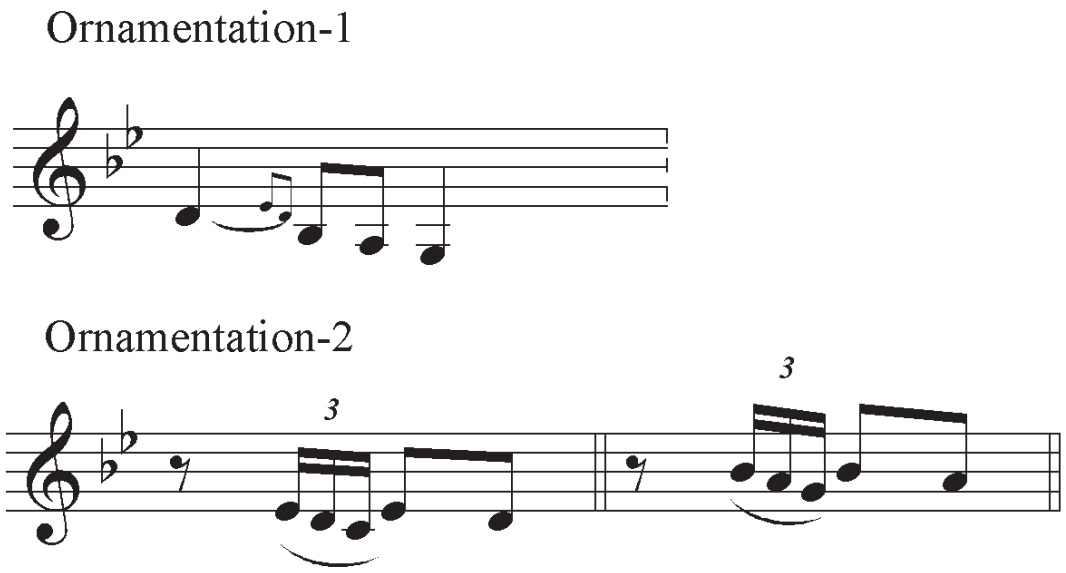


Figure 3: *Shigin* ornamentations.

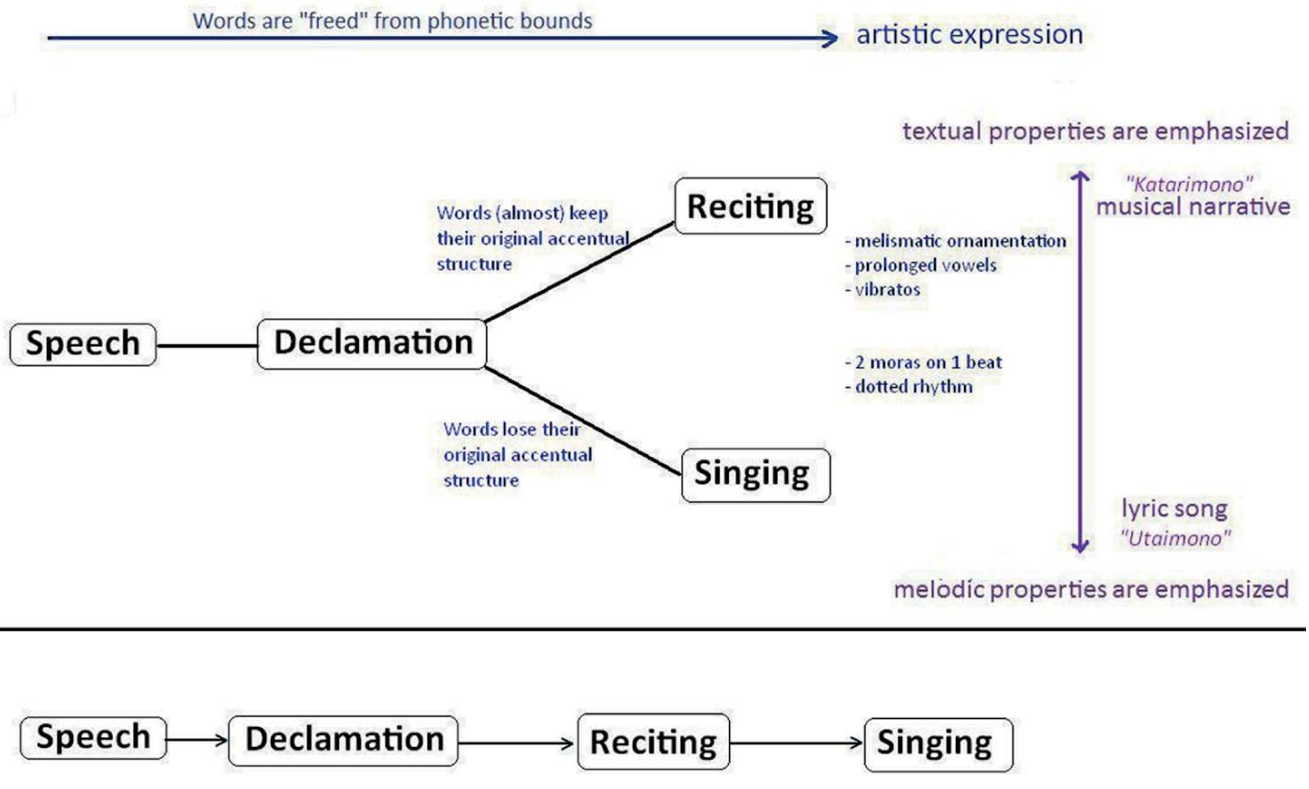


Figure 4: From words to music: conceptual model for the vocal delivery styles of Japanese music.

Please visit http://www.phonogrammarchiv.at/7DE709JKSAG for additional appendices.

References

Cooper, Arthur. 1973. *Li Po and Tu Fu: Poems Selected and Translated with an Introduction and Notes*. London: Penguin books.

De Ferranti, Hugh. 1991. "Composition and improvisation in Satsuma biwa". *Musica Asiatica* 6: 102–127.

Hirano, Kenji. 1990. "Katarimono ni okeru kotoba to ongaku" ['Language and music in performed narratives']. *Nihon Bungaku* 39: 33–43.

Kindaichi, Haruhiko. 1978. *The Japanese Language*. Tokyo: Tuttle.

Kindaichi, Haruhiko. 1989. "Nihon no uta to kotoba" ['Japanese songs and speech']. In: Gamo, S. et al. (eds.). *Iwanami kôza Nihon no ongaku, Ajia no ongaku.* Vol. 7. Tokyo: Ongaku no tomosha, 283–304.

Koizumi, Fumio. 1980. "Rhythm in Japanese Folk Music". In: Emmert, Richard, Yuki Minegishi & Kokusai Kōryū Kikin (eds.). *Musical Voices of Asia: Report of Asian Traditional Performing Arts 1978*. Tokyo: Heibon sha, 108–119.

Koizumi, Fumio. 1984. *Nihon dentô ongaku no kenkyû* ['Studies in Japanese traditional music']. *II: Rhythm*. Tokyô: Ongaku no tomosha.

Koizumi, Fumio. 1997. *Ongaku no kongen ni arumono* ['The origins of music']. Tokyo: Heibonsha.

Komoda, Haruko. 2008. "Biwa no nagare: Satsuma biwa, chikuzen biwa, gendai e". In: Kojima, K. (ed.). *Nihon dentô ongaku kôza "Ongaku"*. Tokyo: Tankô sha, 213–433.

Sato, Yoshiaki.1999. *J-POP Shinkaron* ['The Evolutionism of J-Pop']. Tokyo: Heibonsha.

Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge Language Surveys.

Takeuchi, Lone.1999. *The Structure and History of Japanese*. London: Longman.

Terada, Minoru. 2003. *Seiyashi: 80 Chinese Poetry Classics Roudoku Series*. King Record, KICG 5007, compact disc.

Yamamoto, Kinjô. 1999. *Seiyashi*: *Gin'ei*. Victor, VICG-41084, compact disc.

Yanagi, Kaneko. 1975. *Seiyashi: Gendai nihon kakyoku shû*. Audio Lab Record, OVCA-00001, compact disc.

Yasuko, Sakai. 2004. "The Rhythm at the Time of Singing Japanese: The Prolongation of Child's Spontaneous Songs, Play Songs and Gidayû-Bushi". *Kobe University Department Bulletin* 11 (2): 377–388.

4. Feldforschungsbericht

Die Wayuu in Bild und Ton – Ein Feldforschungserfahrungsbericht

(Barbara Kazianka)

Das Forschungsvorhaben

Im Zentrum meiner Forschung bei den Wayuu, die von Februar bis Juli 2012 realisiert wurde, stand die Frage nach den Verflechtungen von Medizin und Identität. Das Thema generierte sich aus den Daten meiner insgesamt neunmonatigen Feldforschung (in den Jahren 2005 und 2006) bei den Itzá Maya im Norden Guatemalas. Diese diskutierten anhand traditioneller Medizin, was es bedeutete, Itzá zu sein. Im Rahmen meines Doktorats stelle ich meine Beobachtungen zu Beziehungen zwischen traditioneller Medizin und Identitätsdiskursen in den Mittelpunkt. Die konkrete Forschungsfrage lautet, welche Rolle traditioneller Medizin in der Konstruktion von indigenen Identitäten zukommt. Dabei soll der Vergleich zwischen den Itzá und den Wayuu im Umgang mit Medizin und Identität den Nachweis ermöglichen, dass der medizinische Diskurs einerseits Raum, andererseits Objekt von Identitätsverhandlungen ist. Des Weiteren wird der Frage nachgegangen, wie Medizin als primäre Ressource zur Erreichung eines sekundären Zieles im Sinne der Theorie von Paul Unschuld (1975) instrumentalisiert wird. Dieses Ziel ist hierbei nicht Heilung, sondern die Konstruktion einer bestimmten Identität in einem jeweils ganz bestimmten Setting (Zeit, Raum, Machtgefüge, teilnehmende Personen, …), um an eine andere, sekundäre Ressource (sozialer Status, Macht, Fördergelder, …) zu gelangen.

Im vorliegenden Beitrag werden meine Erlebnisse der Datenerhebung mittels Video- und Audioaufnahmen bei den Wayuu beschrieben. Ich möchte an dieser Stelle betonen, dass dieser Bericht eine subjektive Sicht auf Ereignisse, Phänomene wiedergibt und somit ein Produkt von (Selbst) Reflexionen ist. Meine Arbeitsweise bei den Wayuu unterschied sich wesentlich von jener, die ich bei den Itzá anwenden konnte. Teilweise nehme ich auf die Divergenz im Umgang mit Video- und Audioaufnahmen Bezug.

Die Wayuu – eine Einführung

Die Wayuu sind eine indigene Gruppe, deren Habitat die im Norden Kolumbiens und Venezuelas liegende Halbinsel La Guajira ist. Sie werden auf ungefähr eine Million Menschen geschätzt. Ihre Sprache ist das Wayuunaiki,

dass der Arawak-Sprachfamilie zugerechnet wird. Sie leben in einer stark hierarchischen, stratifizierten Gesellschaft, die sich auf ein matrilineares, polygynes Sozialsystem stützt. Ihr Verwandtschaftssystem basiert auf so genannten Klans oder Kasten. Jeder Klan wird auf gemeinsame Ahnen und ein gemeinsames Territorium zurückgeführt. Die Größe eines Klans variiert stark und kann von ein paar hundert bis zu mehreren tausend Angehörigen zählen. Die Wayuu unterscheiden daher zwischen Familie aufgrund von jüngster Blutsverwandtschaft und jener Familie aufgrund gemeinsamer Ahnen. Die Familie bzw. der Klan des Vaters wird als „schwach" verwandt interpretiert und spielt eine untergeordnete Rolle. Dennoch ist sie wichtiger Allianzpartner bei sozialen Ereignissen wie Konfliktbewältigung unter Klans oder Brautpreisbeschaffung. Jede Familie besitzt einen eigenen Familienfriedhof, der wesentlich für die eigene Identität ist.

Aufgrund des ariden Klimas und der Abgeschiedenheit des Großteils der Region lebt die Mehrheit von Subsistenzwirtschaft. Wichtigster symbolischer und ökonomischer Wert sind Viehherden. Damit ihre Tiere in der Wüste und Steppe überleben können, ziehen die Wayuu innerhalb der Guajira, die in Baja (Unter-), Media (Mittel-) und Alta (Ober-) Guajira eingeteilt wird, von einem Familienbesitz zum anderen. Neben Viehzucht, Fischfang, Salzabbau und -verkauf verdingen sich viele männliche Wayuu als Tagelöhner. Hauptsächlich sind die Regionen der Media und der Alta Guajira von Landflucht in städtische Gebiete betroffen. Durch den rechtlichen Sonderstatus, den die Wayuu als indigene Ethnie von Seiten beider Nationalstaaten genießen, sowie aufgrund der geopolitisch strategischen Position der Guajira bestreiten viele ihren Lebensunterhalt durch irreguläre Export-Import Geschäfte. Des Weiteren erkennen die Nationalstaaten die Rechtsprechung der Wayuu an, sodass beide Rechtssysteme, nationalstaatliche und indigene, nebeneinander existieren, was gelegentlich zu Konflikten führt. Sofern die Wayuufrauen keine Berufsausbildung haben, bestreiten sie ihren Lebensunterhalt durch die Produktion und den Verkauf von traditionellen Handarbeiten wie Taschen, Schuhen, Hängematten und Ähnlichem.

Die Wayuu teilen Klans und Familien in ein Klassifikationssystem ein, dass unserem System von Unterschicht, Mittelschicht und Oberschicht ähnelt. Kriterien für diese Einteilung sind Anzahl der Angehörigen des eigenen Klans (dies wird mit Einfluss und Kampfstärke bei bewaffneten Konflikten gleichgesetzt), Reichtum im Sinne von materiellen Gütern wie Vieh (Ziegen, Kühe, Schafe, Pferde, Mulis), Land, Geld sowie anderen Gütern für den Handel wie Salz, symbolische/ökonomische Artefakte wie die *Kakuuna* (Halsketten aus bestimmten seltenen roten Korallen) und Reichtum der Familie an symbolischen/ideellen Gütern wie Anzahl der *Palapreros* (männliche Vermittler bei Konflikten), Anzahl von indigenen

FührerInnen, Ruhm durch korrektes, ehrbares Verhalten und Wissen um die Traditionen, Werte und Normen, etc.

Jene Familien, die in die Mittelschicht bzw. Oberschicht gehören, besitzen neben einem Haus im urbanen Gebiet eine *Rancheria*. Die *Rancheria* ist die traditionelle Siedlungsform der Wayuu. Sie besteht aus einem Haupthaus, der *Loma* (einem überdachten Platz, an dem Besucher empfangen und Gäste in Hängematten übernachten können), einem Kochhaus, einem *Coral* (abgezäuntes Areal, in dem die Ziegen und Schafe sowie teilweise Kälber über Nacht zum Schutz vor wilden Tieren und Viehdieben untergebracht werden) und einer Stelle zum Waschen. Sanitäre Anlagen wie Toiletten gibt es nur in städtischen Gebieten oder bei finanzkräftigen Wayuu. Strom wird mit Hilfe von Generatoren, die mit Öl oder mit Benzin betrieben werden, erzeugt. Einen Wasserbrunnen besitzt ebenfalls nicht jeder. In den meisten Gebieten teilen sich viele Familien ein Wasserloch und tragen das Wasser kanisterweise oft mehrere Kilometer bis zu ihrer *Rancheria*.

Dagegen gibt es in urbanen Gebieten jeglichen finanzierbaren Komfort wie Klimaanlagen. Mobiltelefone und Internet (hier hauptsächlich über Smarttelefone) sind für jeden Wayuu, der es sich leisten kann, ein Muss. Viele soziale Kontakte funktionieren über Kilometer hinweg durch Facebook. Fast jede/r AlphabetIn hat einen Facebook-Account, den er/sie je nach Internetzugangsmöglichkeit nutzt. Offenkundig leben die Wayuu in zwei Welten, die sie geschickt miteinander verbinden: der okzidentalen und der des „Nordens der Wayuu".

Die Aufnahmen im Kontext: Feldforschung – ein schwieriger Prozess

Die anthropologische Feldforschung verlangt dem/der FeldforscherIn sowie den – salopp formuliert – „Erforschten" viel ab. Der/die FeldforscherIn soll – metaphorisch gesprochen – in die Welt der „zu Erforschenden" eintauchen und nach einem Bummel unter Wasser in seiner/ihrer Welt auftauchen, um die Unterwasserwelt der Oberwasserwelt zu erklären. Dieses Unterfangen ist jedoch keineswegs einfach und setzt voraus, dass die BewohnerInnen jener Unterwasserwelt bereit sind, jemanden eintauchen zu lassen. Die Wayuu sind dazu gerechtfertigterweise nur bedingt bereit. Sie bestimmen die Tiefe und die Weite des Tauchgangs. Die Wayuu überleben seit Jahrhunderten in einer relativ unwirtlichen Umwelt und trotzten allen Eroberern. Sie wurden nie besiegt. Einer der Gründe ihres Überlebens liegt in ihrem Stillschweigen, ihren Geheimnissen. Diese hüten sie nicht nur vor Außenstehenden, sondern auch vor anderen Wayuu. Hiervon ist der Bereich der traditionellen Medizin besonders betroffen. Soweit mir bekannt ist, haben nur sehr wenige AnthropologInnen es geschafft, am

Grund entlang zu schnorcheln. Einer von ihnen ist Michel Perrin, der zehn Jahre bei ihnen lebte und ihre Sprache perfekt beherrschte. Soviel Zeit hatte ich nicht zur Verfügung. Meine Feldforschung stellte daher eine ungeahnte Herausforderung dar.

Ich wohnte in Manaure bei einer angesehenen Familie, die politisch sehr aktiv ist. Manaure ist ein Dorf, das für seinen Salzabbau bekannt ist. Es verfügt über befestigte Straßen, mehrere Schulen, ein Krankenhaus, nationale Gesundheitseinrichtungen (EPS und IPS), Apotheken, Internetcafés, verschiedenste Krämerläden und Strom. Das Haus von Rosario „Chayo" Epieyu wurde zu meiner Basisstation. Ihr Enkel Darwin Manuel Cotes Pushaina wurde mein Wegbegleiter während meiner Forschung. Er war Freund, Ratgeber, Führer und Übersetzer. Er verlangte für seine Dienste keinen Lohn, jedoch mussten seine Ausgaben beglichen werden. Er stellte viele Kontakte her und übersetzte mir, wenn eine Person nur Wayuunaiki sprach bzw. ein Thema es erforderte, in Wayuunaiki zu sprechen. Kurz gesagt: Er, seine Großmutter Rosario „Chayo" Epieyu und seine Tante Aleida Alverado Epieyu, die mit mir im Haus ihrer Mutter samt ihren vier Kindern lebte, waren das, was gemeinhin als SchlüsselinformantInnen bezeichnet wird. Mein soziales und forschendes Netzwerk erweiterte sich aufgrund der Mithilfe der Familie Epieyu. An dieser Stelle möchte ich mich nochmals dafür bedanken.

Im Laufe meines Aufenthaltes bereiste ich fast die gesamte Alta und Media Guajira, um die Unterschiede der Lebensbedingungen und Lebensweise kennen zu lernen. Das Datenmaterial, das ich sammelte, waren Fotos, Zeitungen, Bücher, Dokumentationen, private Erzählungen sowie Geschichten und offizielle Interviews, wobei letztere den kleinsten Teil ausmachten. Video wurde mit Sony DCR-SX 65 und Audio mit Edirol R09HR + ORTF-Mikrophon-Aufstellung aufgenommen.

Die Aufnahmebedingungen und -situationen

Der Unterschied zu meinen Erfahrungen mit den Itzá in Aufnahmesituationen war markant. Diese wollten immer alleine interviewt werden. Es machte ihnen nichts aus, aufgenommen zu werden; sie wollten jedoch anonymisiert werden und erhielten ein minimales Entgelt. Die Wayuu verlangten normalerweise kein Geld, jedoch brachte ich Gastgeschenke mit oder kaufte ihnen Waren ab. Die Aufnahmesituationen in der Guajira waren sehr unterschiedlich. Die Mitglieder meiner Gastfamilie konnte ich im Haus in Manaure interviewen. Hier war es möglich, zumindest teilweise auf meinen Schlafraum auszuweichen, wenn die Kinder zu laut spielten oder einfach Persönliches besprochen werden sollte. Alle anderen Interviews fanden in den Häusern meiner InformantInnen statt. Dort bestimmten meine

GesprächspartnerInnen die Umstände der Interviews wie Zeit, genauer Ort, allein oder mit Familie, etc. Da es eine Tradition der Zusammenkünfte gibt, um in Gesprächen Wissen über die Kultur weiterzugeben, war es oft unmöglich, nur mit einer Person zu sprechen, ohne dass die ganze im selben Haushalt wohnende Verwandtschaft zuhörte oder ihre eigenen Interpretationen mitteilte. Auf der einen Seite entstanden dadurch spontane Gruppendiskussionen, auf der anderen Seite konnte allzu Persönliches, das die anderen Familienmitglieder nicht wissen sollten, nicht mitgeteilt werden. Infolgedessen verlor ich einerseits Wissen durch konkrete Beispiele, andererseits gewann ich Informationen über den allgemeinen Konsens und die öffentliche Diskussion zu bestimmten Themen. Aufgrund dieser Umstände wusste ich nie im Voraus, wie das Interview ablaufen würde. Ich bediente alleine alle Aufnahmegeräte und Fotokameras, wodurch ich gezwungen wurde, die Aufnahmegeräte einfach laufen zu lassen, ohne mich viel um sie kümmern zu können. Damit mir keine Informationen durch etwaige Unterbrechungen, Störungen oder Ausfälle entgingen und ich Zeit hatte, mich auf mein Gegenüber zu konzentrieren, verwendete ich meistens Audio- und Videoaufnahmen gleichzeitig. Durchschnittlich dauerte ein aufgenommenes Gespräch zwei Stunden.

Medizinische Praktiken aus dem professionellen Bereich der HeilerInnen der Wayuu (*Outsü*, oder *Piachi*) konnte ich nur bedingt aufnehmen, da im eigentlichen Ritual nicht einmal die eigene Familie zugegen sein darf. Um zumindest einen Einblick in diese Praktiken zu bekommen, musste ich mich selbst behandeln lassen. Diese Aufnahmen besitzen einen simulativen und aktuellen Charakter: simulativ, weil Wasinu Ipuana, die sich bereit erklärte, an mir kleinere Reinigungen durchzuführen, aufgrund ihres Alters eigentlich nicht mehr praktizierte; und aktuell, da Visionen, Träume etc., die im Ritual passieren, nicht vorhersehbar waren.

Befragungen der Orakel (*Ourakuy*), wie jene von Juliana Perez Epieyu, sind öffentlich zugänglich. Das Aufnehmen dieser Sitzungen ist gegen ein Entgelt, das der Geist, der von dem Medium Besitz ergreift, bestimmt, problemlos möglich. Insgesamt konnte ich zwei solcher Orakelbefragungen in Bild und Ton festhalten.

Während meiner Datenerhebung stellte ich ein interessantes, wenn auch nicht ungewöhnliches Phänomen fest: Es war nicht leicht, von Wayuu das Einverständnis zu erhalten, unsere Gespräche aufzunehmen. Zu groß war die Angst vor Missbrauch. Viele Informationen über alltägliche Dinge, Ereignisse, Bräuche, etc., die ich mit meinen Freunden besprach, wurden mir bereitwillig im informellen Raum mitgeteilt bzw. deren tiefere Bedeutung erklärt. Wenn ich jedoch in den offiziellen Interviews darum bat, bereits Bekanntes zu wiederholen, damit ich es aufnehmen konnte,

wichen viele meiner GesprächspartnerInnen aus, sprachen nur allgemein oder weigerten sich rundheraus. Meine Gastmutter Aleida Alverado Epieyu brachte es in einem Interview auf den Punkt: Sie hatte Angst davor, dass die „Geheimnisse der Wayuu" gelüftet und Stereotypen gebildet würden. Konkret wollte sie nicht, dass man die spirituellen Bräuche mit Hexerei in Verbindung bringt und – ihrer Ansicht nach – damit ein negatives Bild über die Wayuu erzeugt. Es wurde scheinbar angenommen, dass alles auf Audio- und Video Aufgenommene öffentlich gezeigt wird bzw. in meine Forschung einfließt, während all jene Informationen, die man mir in Alltagsgesprächen mitteilte und die ich in mein Feldtagebuch schrieb, nicht als Teil meiner Forschungsdaten empfunden wurden.

Aufgrund dieser Ängste waren die Reaktionen der Wayuu auf meine Bitte, unsere Gespräche aufzuzeichnen und die Aufnahmen im Phonogrammarchiv deponieren zu dürfen, sehr unterschiedlich. Manche meiner GesprächspartnerInnen wollten nicht aufgenommen werden. Andere gestatteten mir zwar, unsere Gespräche digital festzuhalten, jedoch unter der Bedingung, dass nur ich die Aufnahmen sah. Wieder andere wollten die Aufnahmen mit dem Vermerk von Sperren archiviert wissen, wenn persönliche Geschichten enthalten waren. Und dann gab es einige Gespräche, die meiner Einschätzung nach zu persönlich waren, um sie archivieren zu können. Infolgedessen stellen die im Phonogrammarchiv zugänglichen Aufnahmen nur einen kleinen Teil meines gesamten Datenmaterials da.

Die Benutzung der Aufnahmegeräte durch die Wayuu

Feldforschungen sind von einer Reziprozität geprägt. So wie ich meine Wayuu-Freunde bat, mir bei meiner Forschung behilflich zu sein, so wollte ich ihnen helfen, wo ich konnte. Eine dieser Möglichkeiten ergab sich durch mein Equipment. Überall dort, wo man filmen oder Fotos machen wollte, die Geräte aber fehlten, wurde ich zu Dokumentationen eingeladen. Ich konnte als Gast Begräbnisse besuchen und dokumentieren, wenn ich die Fotos den Angehörigen zur Verfügung stellte. Ich wurde gebeten, bei wichtigen politischen Anlässen, wie dem Besuch des damaligen Innenministers von Kolumbien, Germán Vargas Lleras, oder der Diskussion über die Verwendung von öffentlichen Geldern in Millionenhöhe zur Entwicklung der Region Guajira, zu filmen. Bei einem Zusammentreffen der indigenen Organisation Outkajuwaa sau'u Waikuaipa nahm ich die Präsentation von Ramon Fernández auf. Er unterbreitete seine Theorien über die Eroberung der Guajira und der Geschichte der Wayuu, die teilweise der offiziellen Geschichtsschreibung widerspricht. Da er diese Theorien bisher noch nicht publiziert hatte, traten einige Wayuu an mich heran, ihnen dieses Material zur Verfügung zu stellen. Ein anderes Mal wurde ich gebeten, die

Beweisfotos am Tatort eines Diebstahls zu schießen. Die Wayuu borgten sich auch gelegentlich meine Fotokameras aus, um Familienfeste und andere Aktivitäten festzuhalten. Der Besitz von Aufnahmegeräten und Fotokameras ermöglichte mir Zugang zu sozialen Aktivitäten, die mir sonst verborgen geblieben wären. Mein Equipment und ich wurden auch benutzt, um bestimmte (politische) Ziele zu verfolgen. Ein deutliches Beispiel hierfür ist meine Aufnahme in Machiques, einem kleinen multiethnischen Dorf in Venezuela. Dort existiert aufgrund von Arbeitsmigration eine große Wayuu-Population. Das Gebiet grenzt an das Territorium der Ethnie Yukpa an. Nun wurde anscheinend in einigen Lokalzeitungen berichtet, dass es Spannungen zwischen diesen beiden Gruppen gebe, da Wayuu in die Gebiete der Yukpa eindrangen. Diese Problematik war mir zum Zeitpunkt meines Aufenthalts nicht bekannt. Als ich ein Interview mit der Beauftragten für indigene Belange, Lilia Jusayuu, halten wollte, entwickelte sich daraus eine Gruppendiskussion mit anderen politisch aktiven Wayuu. Diese bestanden darauf, Juan Mendeola als Vertreter der Yukpa einzuladen und ihn über seine Meinung zu diesem Konflikt zu befragen. Ich empfand die Situation skurril, da Juan praktisch erpresst wurde, zu erscheinen und mir von seinen Erfahrungen zu berichten. Obwohl alle wussten, dass ich Anthropologin und keine Journalistin bin, wollten sie unbedingt vor der Kamera diesen regionalen Diskurs erörtern und klarstellen, dass Wayuu und Yukpa friedlich ohne Probleme koexistieren. Ich bin überzeugt, dass nur deshalb so viel Druck ausgeübt wurde, dieses Thema zu erörtern, da ich unser Gespräch aufnahm. Sie benutzten mich und das Equipment, um ihre politische Botschaft aufzunehmen und öffentlich bekannt zu machen.

Eine weitere Verwendung von Kameras und Mikros findet sich in der Selbstdarstellung auf Medien wie YouTube und Facebook. Neben Fotos werden auch Audio- und Videoaufnahmen ins Netz gestellt. Einige Wayuu filmen Dokumentationen und Filme über ihre Kultur und Lebensweise. Vor allem das Medium Video wird als Repräsentationsmöglichkeit wahrgenommen.

Fazit

Wie in diesem Bericht deutlich wird, ist der Umgang der Wayuu mit Audio- und Videoaufnahmen sehr ambivalent. Auf der einen Seite werden sie als Gefahr gesehen, Geheimnisse einer großen Öffentlichkeit zugänglich zu machen oder negative Stereotype zu kreieren, auf der anderen Seite werden sie verwendet, um das eigene Bild nach ihren Wünschen zu formen oder Botschaften zu verbreiten. Die Manipulationsmöglichkeit von Video- und Audioaufnahmen wurde von den Wayuu erkannt und wird bewusst eingesetzt oder gemieden. Dies gilt insbesondere für selbst gedrehte Videos

der Wayuu. Teilweise hilft der Besitz von Aufnahmegeräten dem/der FeldforscherIn, Kontakte zu knüpfen, und er erleichtert bzw. ermöglicht den Zugang zu bestimmten sozialen Aktivitäten, auf der anderen Seite kann die Tatsache, dass Gespräche, Ereignisse etc. aufgenommen werden, auch dazu führen, dass diese an Qualität verlieren, da die Beteiligten nur mehr Informationen preisgeben, die sie als völlig unbedenklich empfinden. Das heißt, die Forschung kann an Tiefe verlieren. Hier ist der/die FeldforscherIn gefordert, die Situation im Kontext zu sehen und abzuwägen. Manchmal ist es eben sinnvoller, auf Aufnahmen zu verzichten und gute Informationen zu memorieren und/oder per Hand zu protokollieren. Abschließend ist zu erinnern, dass die gemachten Aufnahmen nicht eine Ethnie repräsentieren können, sondern selektive kleine Momentaufnahmen darstellen, die maximal einen Eindruck von Lebenswelten vermitteln können.

Literaturverzeichnis

Perrin, Michel. 1995. *Los practicantes del sueño. EL chamanismo wayuu.* Venezuela: Monte Àvila Editore Latinoamericana.

Unschuld, Paul U. 1975. “Medico-cultural conflicts in Asian settings: an explanatory theory”. *Social Science and Medicine* 9: 303–312.

5. Rezensionen

Margit Berner, Anette Hoffmann & Britta Lange. 2011. *Sensible Sammlungen. Aus dem anthropologischen Depot.* (Fundus-Bücher, 210; ilinx-Kollaborationen). Hamburg: Philo Fine Arts.

Das Ziel dieser „ilinx"-Reihe ist es, Kollaborationen „nicht als neue Ideologie kollektiver Autorschaft, sondern als Bekenntnis zu diskursiven und kooperativen Praktiken" (S. 277) zu sehen. Die Zusammenarbeit der Autorinnen zeigt dies auf überaus bemerkenswerte Weise anhand eines sehr problematischen Themas.

Dieser Band ist eine kritische Auseinandersetzung mit der akademischen Sammlungstätigkeit und somit mit der Geschichte der Anthropologie/n. Die Autorinnen nehmen Bezug auf die Arbeiten des Historikers Andrew Zimmermann (2001), der einen Richtungswechsel von der anfangs liberalen, kosmopolitischen Anthropologie im deutschsprachigen Raum zu einer antihumanistischen, antiliberalen, menschenverachtenden und letztlich auch die Verbrechen gegen Menschen während der NS-Zeit antizipierenden Disziplin aufzeigte und diesen Bruch rund um den Ersten Weltkrieg zeitlich festsetzte (S. 21f.).

Die drei Autorinnen nähern sich in zehn Beiträgen auf unterschiedliche Weise ‚sensiblen Sammlungen', die bis in die heutige Zeit in diversen Museumsdepots und wissenschaftlichen Archiven (wie dem Berliner Phonogramm-Archiv, dem Lautarchiv in Berlin, dem Naturhistorischen Museum in Wien und dem Phonogrammarchiv der ÖAW) der kritischen Aufarbeitung harren. Unter ‚sensiblen Sammlungen' werden ‚kulturell sensible Gegenstände' und „menschliche Überreste oder Gegenstände von religiöser Bedeutung" (*ICOM Code of Ethics for Museums* 2004) verstanden, die unrechtmäßig entwendet oder durch ungleichen Handel, Kauf und Tausch erworben wurden. „In diesem Buch geht es um die Geschichten, die solche Objekte erst zu sensiblen Sammlungsbeständen machten. Sensibel sind nicht nur heutige Umgangsweisen mit den Dingen und ihren Veröffentlichungen, sondern auch ihre Provenienz, ihr Transfer, ihre Zirkulation, ihre Herauslösung aus lebensweltlichen Zusammenhängen und letztlich ihre Verwandlung in Sammlungsgegenstände" (S. 19).

Der Band versteht sich als Werkstattbericht. Anhand einzelner Beispiele werden Sammlungsstrategien und Bestände in physisch-anthropologischen Depots und Tonarchiven in Deutschland und in Österreich, von der zweiten Hälfte des 19. bis zur Mitte des 20. Jahrhunderts, analysiert. Die Forscherinnen

recherchieren anhand von Einzelfällen die Entstehungskontexte, die ideologischen Grundlagen der Sammlungs- und Präsentationsstrategien und den heutigen Umgang mit diesen ‚sensiblen Sammlungen'. Trotz mehrerer internationaler und museumseigener Richtlinien bezüglich der „Human Remains Policies" bleibt der institutionelle Umgang damit sensibel. Eine Voraussetzung für einen ethisch verantwortungsvollen Umgang damit ist eine kritische Aufarbeitung der Entstehungskontexte, wie koloniale Machthaber sich die materielle Welt aneigneten (S. 20), und ein diskursiver Umgang, der anhand von Einzelbeispielen facettenreich und deutlich aufgezeigt wird.

Die Autorinnen setzen sich kritisch mit der ‚salvage anthropology' jener Zeit auseinander, deren Ziel es war, ‚aussterbende Völker' durch das Anlegen von physisch-anthropologischen Sammlungen und solchen materieller Kultur, Fotografien, sowie Film- und Tonaufnahmen zu dokumentieren. Es handelte sich um Ethnien, die vor allem aufgrund kolonialer Machtbestrebungen in ihrem Lebensraum und ihrer Lebensweise existenziell bedroht waren. Waren es anfangs noch umfassendere ethnografische Arbeiten, so standen dann oft nur mehr anthropometrische Untersuchungen, Fotografien und die Abnahme von Körperformen in Gips im Mittelpunkt des politisch motivierten Interesses. „Solche Abnahmen entstanden meist in prekären Situationen, die von der Überlegenheit und Definitionsmacht der Wissenschaftler und der Kolonialbeamten geprägt waren" (S. 32).

Unter den *Anleitungen für ethnographische Beobachtungen und Sammlungen*, die „zivilisierte Wissenschafter" jener Zeit propagiert haben, waren genaue Anweisungen, wie ‚anthropologisches Material' – also Skelette – präpariert, wie Gipsabgüsse abgenommen und wie Skalps und Ziernarben gesammelt werden sollen. Selten wird in den wissenschaftshistorischen Arbeiten genau darauf eingegangen, wie große physisch-anthropologische Sammlungen zustande kamen – diese Publikation scheut sich nicht darauf einzugehen, dass Gebeine aus einem Grabraub oder von zu Tode Verurteilten und von Ermordeten aus den KZs stammten. Diese ‚aggressiven Akte des Sammelns' bilden den Ausgangspunkt der Diskussion einzelner Fallbeispiele. Der Fokus einerseits auf physisch-anthropologische und andererseits auf Audio-Sammlungen scheint anfangs ein sehr unterschiedlicher Zugang zu sein, ermöglicht in dieser Zusammenschau aber nicht alleine eine außerordentlich kritische Auseinandersetzung, sondern einen besonderen Einblick in die Machtverhältnisse des Sammelns aus der Sicht der Beforschten.

Bereits im Einleitungskapitel werden, fast Forschungsskizzen gemäß, unterschiedliche ‚andere' Perspektiven auf sensible Sammlungen aufgezeigt und die LeserInnen in den Bann gezogen. Britta Lange führt anfangs in die internationalen rechtlichen Bestimmungen im Umgang mit sensiblen Sammlungen ein und umreißt die disziplinären Grundlagen

sowie ‚wissenschaftlichen' Strategien ab dem Ende des 19. Jahrhunderts. Margit Berner geht in dem Beitrag ‚Schauen und Wissen' auf die ersten musealen Präsentationen der physischen Anthropologie in Österreich ein und zeigt den Konnex zwischen den Völkerschauen, der Zur-Schau-Stellung menschlicher Exponate seit dem Ende des 18. Jahrhunderts und den physisch-anthropologischen Sammlungen im 19. und 20. Jahrhundert, die Schädel, Skelette, Körperteile, Modelle, Fotografien und Gipsabdrücke umfassten. Anette Hoffmanns Artikel über ‚Glaubwürdige Inszenierungen' geht am Beispiel der Biografie von Hans Lichtenecker auf die aggressiven Akte des Sammelns in den 1930er-Jahren im damaligen Südwest-Afrika ein. Die Vermessungen und Gipsabdrücke von menschlichen Körperteilen in einer post- genozidalen Gesellschaft verwirklichte Lichtenecker mittels autoritärer Selbstinszenierung in einer Polizeistation. „Die unter anderem von Alan Sekula aufgezeigte historische Verbindung zwischen Kriminologie, anthropometrischen Vermessungen und Fotografien nimmt hier eine nahezu gespenstische Mimesis an" (S. 82f.). Margit Berners Spurensuche nach den Hintergründen über die Herstellung einer Maske von Lichtenecker und einer abgeformten Ganzkörperfigur zeigt nicht nur die Entstehungskontexte weiter auf, sondern die Strategien und Kooperationen der Museen, europäische Sichtweisen auf ‚die Anderen' zu konstruieren. In Britta Langes Beitrag über Tonaufnahmen von deutschen und österreichischen Wissenschaftlern in Kriegsgefangenenlagern des Ersten Weltkriegs werden deren akademische Verflechtungen und Finanziers sowie die unterschiedlichen sprachwissenschaftlichen Herangehensweisen diskutiert. Sie wirft in diesem Artikel zwei Themenbereiche auf, die sich in den weiteren Beiträgen fortsetzen – die Situation der beforschten Menschen und deren geringe Möglichkeiten, Widerstand gegen und Kritik an dieser Behandlung zu leisten, wie auch das heutige Interesse in den Herkunftsländern der beforschten Personen an diesen Audioaufnahmen. Die angefertigten Tondokumente fanden in der Zeit der Herstellung wenig Beachtung, manche wurden erst in den letzten zwei Jahrzehnten erstmals übersetzt und veröffentlicht. Ausgehend von Lichteneckers Tondokumenten zeigt Anette Hoffmann Strategien des Widerstands auf. Bezugnehmend auf Glissants Konzept der *opacity* (1997) werden unverständliche Sprechakte als strategische Form der Kritik verstanden (S. 132). Genaue Analysen von traditionellen Erzählungen zeigen Versuche, darin versteckt einen Protest an den kolonialen Verhältnissen zu formulieren. Die Autorinnen gehen davon aus, dass die Inhalte der Erzählungen und die Widerstandsformen in der Zeit der Aufnahme von den Europäern nicht verstanden wurden. In einem weiteren Artikel analysiert Hoffmann erstmals 80 Jahre nach der Entstehung der Aufnahmen die Inhalte und Sprechakte der entmachteten

kolonialisierten Bevölkerung. Außerordentlich ist auch, dass durch diese Herangehensweise erkannt wurde, wie in manchen der Tonaufnahmen die Sprecher auf die erfahrenen Gewaltakte der physisch-anthropologischen Untersuchungen Bezug nehmen. „Insgesamt haben die bisher unbearbeiteten Tonaufnahmen das Potenzial, mit ihren Versionen und Interventionen die sonore Stimme der kolonialen Ethnografie zu stören." (S. 184). Margit Berners Beitrag ‚Die haben uns behandelt wie Gegenstände' zeigt, wie die physisch-anthropologischen Techniken und Arbeitsweisen, die man in den Kriegsgefangenenlagern des Ersten Weltkrieges angewandt hatte, von der nächsten Forschergeneration in den nationalsozialistischen Internierungslagern fortgeführt wurden. Eine ‚anthropologische Kommission' unternahm Untersuchungen an jüdischen Männern, die im Wiener Stadion 1939 interniert wurden. Zu Objekten der rassistischen Forschung degradiert, wurden sie ins KZ Buchenwald deportiert, nur wenige überlebten. Margit Berner erarbeitete im Zuge ihres Forschungsprojektes das Schicksal dieser Personen und konnte mit zwei der Überlebenden Kontakt aufnehmen und die Sicht der Opfer darstellen. Der Sammelband endet mit einer kritischen Reflexion von Britta Lange über einen entsprechenden Umgang mit ‚sensiblen' Objekten in physisch-anthropologischen Sammlungen und auch ‚sensiblen' Tonaufnahmen in Schallarchiven, die zwischenzeitlich selbst zu historischen Objekten geworden sind. Lange verdeutlicht, dass es nicht nur eine Strategie geben kann, dass auch Restitutionen nicht immer möglich sind, dass sich aber die Institutionen nicht ihrer Verantwortung entziehen können. ‚Wege aus den Depots', wie auch der Untertitel besagt, müssen durch deren materielle und virtuelle Öffnung, durch diskursive Praktiken und die Absicht, sich den dunklen Seiten der eigenen Wissenschaftsdisziplin kritisch zu stellen, gekennzeichnet sein.

Britta Lange, Anette Hoffmann und Margit Berner haben hier sehr mutig und ambitioniert neue Wege beschritten, haben ‚wissenschaftliche Choreographien' beleuchtet und nicht allein gezeigt, wie sich diese ad absurdum führten, sondern in welchen Gewaltkontexten ethnografische Forschung zeitweise erfolgt ist.

Das im Format sehr klein gehaltene Buch birgt viel Sprengstoff in sich. Die Autorinnen zeigen mit viel Umsicht (Verwendung von Abbildungen), in kritischer Auseinandersetzung mit der eigenen Disziplingeschichte und durch ihre ambitionierten Forschungen, dass wir noch viel genauer die Entstehung von Wissen hinterfragen und auf die politischen Kontexte von Forschungen und deren Protagonisten Bezug nehmen müssen.

Maria Six-Hohenbalken
Institut für Sozialanthropologie
Österreichische Akademie der Wissenschaften

Jarmila Procházková (ed.). 2012. *As Recorded by the Phonograph: Slovak and Moravian Songs Recorded by Hynek Bím, Leoš Janáček and Františka Kyselková in 1909–1912*. 3 vols. (I. Studies and Reports; II. Transcriptions of Texts; III. CD 1–3, DVD). Brno: Etnologický ústav AV ČR.

Mit der Erfindung des Phonographen beginnt ein neues Zeitalter in der Musikforschung, denn erstmals ist es möglich, ein Klangerlebnis festzuhalten und damit wiederholbar zu machen. Nach der rasanten Weiterentwicklung der Technik führt die vorliegende dreibändige Publikation (Untersuchungen und Berichte, Transkriptionen sowie Hörbeispiele und DVD) in die Anfangszeit und zeigt auf, wie damals Tonaufnahmen, nämlich akustische Belege, als wesentliche Grundlagen für die Volksmusikforschung geschaffen wurden.

Wie jede neue Erfindung spaltete auch der Phonograph die Forscherpersönlichkeiten der Zeit bezüglich Brauchbarkeit und Einsetzbarkeit. Lehnten die einen den Apparat als unbrauchbar und unnötig ab, begrüßten die anderen die Möglichkeit der Wiederholbarkeit und damit Überprüfbarkeit eines Belegs. Ebenso faszinierte, dass neben der Melodie auch die Charakteristik und Akzentuierung der Interpretation festgehalten werden konnten.

Gerda Lechleitner geht in ihrem Beitrag über die frühen Volksmusikaufnahmen – unter dem Motto: Wir sammeln niemals ‚das' Lied, sondern nur eine Version von ihm (nach einem Ausspruch von Constantin Brăiloiu) – auf die Entwicklung der Aufnahmetechnik mit ihren Vorteilen, Schwierigkeiten und Konsequenzen für die Volksmusikforschung ein und fügt in gebotener Kürze die unterschiedlichen Positionen der Forscher an. Jarmila Procházková, die Hauptverantwortliche des Projektes, bezieht sich in ihrem Beitrag auf Janáčeks Aufnahmen und deren zeitgenössische Rezeption. Innerhalb des großen Projektes „Das Volkslied in Österreich", das die Sammlung und Dokumentation aller Nationallieder und -melodien der Kronländer der Donaumonarchie zum Ziel hatte, leitete der berühmte Komponist Leoš Janáček den tschechischen Arbeitsausschuss, zu dem er mehrere Experten einlud. Vor allem die zwischen 1909 und 1912 mit dem Phonographen gemachten Aufnahmen von Hynek Bím und Františka Kyselková bilden die Grundlage dieser Publikation. Janáček untersuchte die Aufnahmen selbst, seine Kommentare und Korrekturen betreffen vor allem die Problematik der neuen Technik und die Brauchbarkeit für eine wissenschaftliche Untersuchung.

Eine Beschreibung der fünf Aufnahmesitzungen des Arbeitsausschusses für Tschechische Nationallieder in Mähren und Schlesien enthält die Lieder

in alphabetischer Reihenfolge sowie detaillierte Beschreibungen der einzelnen Besuche, aber auch Namen und Angaben zu den Gewährspersonen. Es ist spannend nachzulesen, wie lange die einzelnen Regionen besucht und wie die Kontakte zu den Sängerinnen und Sängern hergestellt wurden. Man erfährt viel über die Vorgangsweise und die Aufnahmesituationen, sowohl über Erfolgserlebnisse als auch Probleme, sowie über die Erfahrung mit dem Phonographen. Eine sechste Forschungsfahrt war geplant, wurde aber ohne einen Phonographen durchgeführt. Auf die schriftlichen Belege wird hingewiesen.

Hana Urbancová, Alžbeta Lukáčová und Lucie Uhlíková teilen sich sowohl die Transkriptionen als auch die Kommentare zu den Aufnahmen. Die Angaben zu den Aufnahmesitzungen, den Inhalten, Gewährspersonen, Regionen, den Aufzeichnern und den musikalischen Parametern machen die Hörbeispiele auf den CDs erst zu Grundlagen für die wissenschaftliche Forschung.

Die Zylinder, die in der Brünner Abteilung des Ethnologischen Institutes der Tschechischen Akademie der Wissenschaften aufbewahrt sind, wurden schon in den 1950er-Jahren von Olga Hrabalová auf modernere Tonträger überspielt, um eine komplette Liste der Aufnahmen zu ermöglichen. Ihre Ergebnisse bilden einen Ausgangspunkt des aktuellen Projektes, für das noch einmal alle Zylinder übertragen und digitalisiert wurden. Wie in vielen anderen Projekten begleitete auch hier das Phonogrammarchiv die Übertragung und trug mit wesentlichen Hinweisen und Schulungen zum Gelingen bei. Franz Lechleitner, ehemaliger Cheftechniker und Experte für Wachszylinder, schreibt über die Herausforderungen des Übertragungsvorgangs, während Milan Fügner auf den Umgang mit verletzten und beschädigten Zylindern eingeht. So ist es den Technikern gelungen, auch zerbrochene Walzen abzuspielen. Nach der Übertragung bietet die digitale Technik die Möglichkeit einer nachträglichen „Verbesserung“. Man kennt die Nebengeräusche von historischen Tondokumenten, die einerseits immer präsent sind, aber andererseits auch das Nutzsignal bis zur Unkenntlichkeit überlagern können. So bedient man sich – unter Wahrung des originalen Klanges – gerne der technischen Möglichkeit der Restaurierung. Václav Mach untersuchte im Rahmen eines Projektes an der Technischen Universität in Brünn die technischen Parameter des vorhandenen Signals in Hinblick auf die Grenzen der Verbesserungsmöglichkeiten.

Im zweiten Band folgt nach einer Einführung von Jarmila Procházková zu den Transkriptionen und zur Rezeption der Zylinderaufnahmen Leoš Janáčeks und seiner Kollegen (in tschechischer und englischer Sprache) die inhaltliche Beschreibung der CDs: Die ersten beiden CDs enthalten Aufnahmen aus der Slowakei, die von F. Kyselková in den Bergregionen

von Strážov und Javorníky gemacht wurden sowie jene gemeinsam mit L. Janáček in Brünn (1909) aufgenommenen, transkribiert von Hana Urbancová. Die Aufnahmen von H. Bím und F. Kyselková in der Terchová Region sowie die von L. Janáček in Brünn (1912) übertrug Alžbeta Lukáčová. Die dritte CD beinhaltet die Aufnahmen aus Mähren, wo F. Kyselková und H. Bím in Vnorovy unterwegs waren; diese bearbeitete Lucie Uhlíková. Sämtliche Texte sind gedruckt, die Incipits sogar ins Englische übersetzt. Außerdem gibt es hier detaillierte Angaben zu den einzelnen Aufnahmen.

Die DVD enthält alle Lieder und die handschriftlichen Aufzeichnungen der Forscher und der Forscherin mit den Transkriptionen und Texten in übersichtlicher Form und in alphabetischer Reihenfolge. Hana Urbancová, Alžbeta Lukáčová und Lucie Uhlíková erläutern wichtige Charakteristika der Lieder sowie die augenscheinlichsten Unterschiede der schriftlichen und akustischen Belege in ihren Kommentaren in tschechischer und englischer Sprache. Außerdem sind auf der DVD auch jene Aufnahmen enthalten, die aufgrund ihrer schlechten Tonqualität keinen Platz auf der CD gefunden haben. Diese Aufnahmen dienen ausschließlich einer wissenschaftlichen Untersuchung und sind nicht zum bloßen Anhören gedacht.

Zu guter Letzt wünsche ich historischen Aufnahmen eine Aufarbeitung wie diese und dieser Publikation eine kritische und lebendige Rezeption sowie viele Nachahmungen. Herzliche Gratulation zu diesem Projekt!

Michaela Brodl
Archiv, Österreichisches Volksliedwerk

TÄTIGKEITSBERICHT des Phonogrammarchivs für das Jahr 2012

Die finanzielle Situation der Akademie führte zum Verlust von vier Mitarbeitern und einer drastischen Reduktion des Sachbudgets. Nichtsdestotrotz wurde versucht, dem in der Evaluation von 2011 ausgewiesenen Status eines „nationalen Kompetenzzentrums und einer nationalen Sammelstelle für den Bereich AV-gestützter Feldforschung" gerecht zu werden und auch die Entwicklung auf technischem Gebiet konsequent weiter voranzutreiben. Zur effektiven Einbindung des Phonogrammarchivs in die nationale und internationale Forschungsinfrastruktur wurden Kontakte mit den neuen Infrastruktur-Plattformen CLARIN-AT und DARIAH aufgenommen.

1. Erforschung und audiovisuelle Dokumentation bedrohter Sprachen und Musikkulturen

Feldforschung

Li Huang unternahm gemeinsam mit Rudolf M. Brandl und Bernhard Graf eine Feldforschung in Guangzhou und Linwu, Volksrepublik China (08.08.–21.08.2012) zur Dokumentation der Yuju- und Nuo-Oper.

Christian Huber setzte seine Erforschung des Shumcho, einer bedrohten und bislang nicht dokumentierten tibeto-birmanischen Sprache des District Kinnaur (Himachal Pradesh, Indien), fort. Die 2012 durchgeführten Forschungen machten sich auch die modernen Kommunikationstechnologien zunutze und brachten viele neue Erkenntnisse, vor allem zur Intransivierungs- und Objektskongruenzmorphologie sowie dem tonalen Verhalten gewisser Verben. Zudem wurden aufschlussreiche Kontraste zu Kinnauri und anderen verwandten Sprachen aufgedeckt.

2. Kontextualisierende Forschung

2.1. Kontextualisierung, Auswertung und Bereitstellung von audiovisuellen Quellen

Das von der Stadt Wien subventionierte Projekt „Das Phonogrammarchiv in Wien – Spiegel österreichischer Afrikaforschung: Kontextualisierung von Tonaufnahmen in afrikanischen Sprachen" (Leitung: Clemens Gütl, Laufzeit: April 2011–Oktober 2012) ging dem Phonogrammarchiv durch den Einzeltransfer Clemens Gütls an die Universität Wien mit 01.04.2012 leider verloren.

Nach Fertigstellung der Romani-Transkriptionen und deutschen Übersetzungen von Märchen und anderen Erzählungen durch Christiane Fennesz-Juhasz, Mozes Heinschink und Petra Cech (die Letzteren im Auftrag des Romani Projekts des Fachbereichs Plurilingualismus am *treffpunkt sprachen* der Universität Graz) konnte die zweisprachige Textedition zur Erzählkultur der Kalderaš Anfang September beim Drava-Verlag publiziert werden. Der Textband enthält zahlreiche Märchen, Schwänke und andere Geschichten aus der Oraltradition von Kalderaš aus Österreich sowie diversen Ländern Europas und Lateinamerikas. Zusätzlich geben autobiographische Berichte und Gespräche Einblick in Kultur, Alltag und sich wandelnde Lebensbedingungen, aber auch in Erfahrungen systematischer Verfolgung dieser Roma-Gruppe. Als Quellen für diese Sammlung dienten größtenteils im Phonogrammarchiv archivierte Tonaufnahmen, die zwischen 1965 und 2012 im Zuge von einschlägigen Feldforschungen gemacht wurden. Im ausführlichen Kommentar am Ende des Bandes werden die ErzählerInnen vorgestellt und die Quellen und Herkunft der Texte, Textgattungen und Romani-Varietät sowie Geschichte und Kultur der einzelnen Kalderaš-Gruppen der verschiedenen Länder erläutert.

Gemeinsam mit Mozes Heinschink überarbeitete Christiane Fennesz-Juhasz einen 2010 gehaltenen Konferenzbeitrag zu Selbstzeugnissen von Roma, der – ebenfalls weitestgehend auf entsprechenden Archivbeständen (Lebensgeschichten, Zeitzeugenberichte, [akustische] Briefe, Lieder, etc.) basierend – im Proceedingsband bei der innsbruck university press voraussichtlich 2013 erscheinen wird.

Für die Serie 13 "Recordings from the Caucasian Region 1909 and 1915/16" wurden die Aufnahmen von Robert Lach kontrolliert und transliteriert, wobei Diskrepanzen zwischen seinen Transkriptionen und den tatsächlichen Inhalten der Aufnahmen aufgezeigt werden konnten (diese Erkenntnisse wurden beim 6th International Symposium on Traditional Polyphony in Tbilisi vorgestellt). Der allgemeine Kommentar von Britta Lange wurde von Christian Liebl ins Englische übersetzt. Die Vorbereitung für die Druckvorlage und das Erstellen der Master-CDs ist für 2013 geplant. Für die Edition der noch verbleibenden Aufnahmen russischer Kriegsgefangener aus der Zeit des Ersten Weltkrieges wurde ein Subventionsansuchen an die Magistratsabteilung 7 (Kultur-, Wissenschafts- und Forschungsförderung der Stadt Wien für 2012) unter dem Titel „‚Sensible Sammlungen' – Tonaufnahmen aus der Zeit des 1. Weltkriegs, aufgenommen in Kriegsgefangenenlagern der Österreichisch-Ungarischen Monarchie", eingereicht, das leider abgelehnt wurde. Nun wird für diesen Teil der Sammlung ein Projekt beim Jubiläumsfonds der Österreichischen Nationalbank unter dem Titel „Displaced Voices" eingereicht.

Die Arbeit an der Serie 14 (die Aufnahmen von Hermann Junker, Robert Stigler et al.) wurde weitergeführt, die Master-CD erstellt sowie die wissenschaftlichen Kommentare für das Booklet großteils ins Englische übersetzt.

Im Zuge der Recherchen von Ulla Remmer für die Edition der CD „Indian Recordings 1904/5, 1918, 1927“ (Serie 15) konnten bislang nicht bekannte Texte identifiziert werden; die Transkriptionen wurden 2012 weitgehend fertiggestellt, der Kommentarteil ist in Arbeit.

Über Kontakt mit Cristina Ghirardini ist eine Kooperation mit dem „Istituto Friedrich Schürr“ (Ravenna, Italien) angedacht. Grozdana Marosevic (Institut za etnologiju i folkloristiku, Zagreb, Kroatien) teilte mit, dass große Fortschritte am zweiten Band der kroatischen Aufnahmen erzielt wurden, sodass von Seiten des Phonogrammarchivs die technische Aufbereitung (Master-CDs) in nächster Zukunft in Angriff genommen wird.

2.2. Wissenschaftliche Veranstaltungen, Projekte und Kooperationen

Das 19. Meeting der Study Group on Historical Sources of Traditional Music (International Council for Traditional Music, ICTM) fand in Wien, im Clubraum und im Theatersaal der Österreichischen Akademie der Wissenschaften, vom 06.–10.03.2012 statt. Die Themen der Konferenz lauteten: „Historical sources and contemporary fieldwork in ethnomusicology – relationship, dialogue, mutual benefit“ und „Multidisciplinary approaches to the study of historical sources of traditional music“ (s. Konferenzbericht von Bernd Brabec de Mori im *Jahrbuch des Phonogrammarchivs* 3: 94–100).

Der sehr erfolgreiche „Workshop Relationships of Speech Tone and Music (WRSTM)“ (05.–07.07.2012) brachte 16 erfahrene Linguisten und Ethnomusikologen aus 9 Ländern zusammen. In 13 transdisziplinären Präsentationen wurde dem vielfältigen Verhältnis von Sprachton und Musik in Tonsprachen nachgegangen. Die Anwesenden waren einhellig der Meinung, dass eine stärkere Zusammenarbeit zwischen Linguistik und Ethnomusikologie in diesem Bereich wünschenswert ist.

Die Mitarbeit des Phonogrammarchivs am von der EU mitfinanzierten Projekt „European Acoustic Heritage“ setzte sich im Berichtszeitraum fort und zeitigte Ergebnisse wie den Start der Wanderausstellung (Tampere, September), die Einbeziehung von Beständen des Phonogrammarchivs in die „Soundscape Map“ (<http://map.europeanacousticheritage.eu>), weitere Wiederholungen historischer Aufnahmen sowie Neuaufnahmen in Raumklang-Technik (Culture Programme 2007–2013, Strand 1.2.1, Agreement No. 2011–0521/001–001 CU7–COOP7, Laufzeit 01.05.2011–30.04.2013; J. Schöpf/Projektmanagement, N. Wallaszkovits, C. Fennesz-Juhasz, J. Spitzbart).

Im 2011 angelaufenen EU-Projekt „Introducing interdisciplinarity in music studies in the Western Balkans in line with European perspective – InMuSWB“ (G. Lechleitner/Projektmanagement, C. Fennesz-Juhasz, N. Wallaszkovits, J. Spitzbart; EACEA-TEMPUS IV, Joint Projects; 517098-TEMPUS-1-2011-1-RS-TEMPUS-JPCR; Okt. 2011–Okt. 2014) standen die Teilnahme an der Etablierung des Forums von Musikuniversitäten (in Sarajewo) sowie ein „teacher's training“ (in Wien) und Kurse in Digitalisierung und Archivierung (in Sarajewo) auf dem Programm.

Im Dezember 2012 wurde das von der Fundação calouste Gulbenkian (Portugal) finanzierte Projekt „Recuperação e tratamento documental de uma coleção de discos de 78rpm de música portuguesa doados à Universidade de Aveiro“, eine Kooperation zwischen dem Department für Ethnomusikologie der Universidade Aveiro (Portugal) und dem Phonogrammarchiv, mit einem zweiwöchigen Workshop in Aveiro begonnen.

Zusammen mit Michael Studemund-Halévy (Institut für die Geschichte der deutschen Juden, Hamburg) und Ivana Vučina Simović (Universität Kragujevac, Serbien) arbeitete Christian Liebl an der Herausgabe der Tagungsakten der internationalen Konferenz „Sefarad an der Donau“ (ÖAW, Wien 2011). Der Band wird auch zusätzliche Materialien zur Geschichte der sefardischen Gemeinde Wiens enthalten und 2013 publiziert werden.

Die Kooperation mit dem Ethnologischen Institut der Akademie der Wissenschaften der Tschechischen Republik, v.v.i. (Forschungsstelle Brno und Kabinett für Musikgeschichte), wurde mit der Veröffentlichung von Vzaty do fonografu (As Recorded by the Phonograph: Slovak and Moravian Songs Recorded by Hynek Bím, Leoš Janáček and Františka Kyselková in 1909–1912) zum erfolgreichen Ende gebracht.

In Kooperation mit dem Institut für Volksmusikforschung und Ethnomusikologie der Universität für Musik und darstellende Kunst in Wien wurde im Rahmen von „Urban Europe – Joint Programming Initiative“ das Projekt „Music as a Social Medium in Urban Centres as a Means to Deal with Diversity and Social Cohesion: Laying the Ground for New Approaches in Research and Cultural Policy on Immigration“ eingereicht (weitere Partner Initiative Minderheiten und Svenskt visarkiv); nach positiver Evaluierung in der ersten Runde wurde das Projekt jedoch letztlich nicht bewilligt.

Das CD-Projekt „Klingende Forschung – Konzertdokumentation“ ist eine Kooperation mit der Kommission für Musikforschung (seit 2013: Abteilung Musikwissenschaft des Instituts für kunst- und musikhistorische Forschungen) der ÖAW. 2010 wurden durch das Phonogrammarchiv die ersten (Gesprächs-)Konzerte akustisch dokumentiert; im Projekt ist das Phonogrammarchiv für die Erstellung der Master-CDs sowie für Redaktion und Organisation verantwortlich. 2012 erschienen die ersten drei

CDs (s.u., Publikationen). Die erste CD dieser Reihe (Schubert Adrast – Einzelnummern und Entwürfe zu einer Oper D 137) wurde im September 2012 mit dem Pasticcio-Preis von Radio Ö1 ausgezeichnet.

2.3. Feldforschung

Katharina Thenius-Wilscher setzte nach Beendigung ihrer Elternkarenz das Feldforschungsprojekt zu aktuellen Tendenzen in der Volks- und Populärmusikpraxis in Österreich fort. Im Sinne einer Re-Study konnten weitere Vergleichsaufnahmen zu dem vom Phonogrammarchiv in den Jahren 1978–85 durchgeführten Projekt „Singen und Musizieren in Österreich bei aktuellen Anlässen" hergestellt werden.

Hedwig Köb beendete ihre Feldforschungsreise in Senegal (26.10.2011–04.01.2012) zur Dokumentation der in den frankophonen islamischen Ländern Westafrikas verbreiteten Griot-Tradition im soziokulturellen Kontext ihrer Heimat als Ergänzung zu den einschlägigen stadtethnologischen Erhebungen in Österreich.

Im Forschungsschwerpunkt „Mechanische Musikinstrumente" (Helmut Kowar) wurde die Musik zweier weiterer Renaissanceautomaten des Kunsthistorischen Museums mittels einer Rekonstruktion wieder zum Klingen gebracht. Der komplexe Automat in Form eines Schiffes (1585) mit zwei Musikautomaten und das automatische Regal des Bacchuswagens (ca. 1602–1606) verfügen über gut erhaltene Programmspeicher (Metallplatten) sowie über teilweise noch spielbare Zungen bzw. Lippenpfeifen und eine noch funktionierende Trommelmechanik. Einzeltonaufnahmen aller originalen Quellen bzw. der Nachbauten einiger Pfeifen dienten als Grundlage für eine Rekonstruktion der Musik, die damit weitestgehend dem ursprünglichen Spiel und Klangbild der Automaten entspricht. In gleicher Weise wurde die Tonfolge des dreitönigen automatischen Regals des Glockenturmes (um 1580) rekonstruiert. Mit Ton- und Videoaufnahmen der wieder spielbar gemachten Orgelwerke des Walbaum-Schrankes (um 1620/25) und des Triumphwagen mit Minerva (1625–30) wurden die Dokumentationen der Automaten und die Wiederherstellung ihrer klingenden Musik abgeschlossen.

2.4. Forschungsdatensammlung

Christian Huber schloss die Archivierung der Sammlungen von dokumentarischen Aufnahmen aus Spiti von Sonam Tsering (2002 und 2003) ab. Hedwig Köb beendete die Archivierung der Audioaufnahmen des Projekts „Mbalax im Senegal (Sommer 2005)" (Rupert Krieger, Institut für Musikwissenschaft, Universität Wien) sowie die Archivierung der Videofeldforschungsaufnahmen aus Brasilien 2010 von Barbara Alge.

Die sukzessive Digitalisierung der zwischen 1950 und 1990 auf Papier verfassten Protokolle zu den analogen Archivbeständen (inkl. schriftlicher und fotografischer Materialien) musste aufgrund von Personalkürzungen (Nicht-Verlängerung des Vertrags von Florian Hoidn) hintangestellt werden.

Im Herbst 2012 wurde mit der Digitalisierung der Aufnahmen, die Karoly Gaál zwischen 1961 und 1973 mit Unterstützung des Phonogrammarchivs erstellt hatte, begonnen (C. Fennesz-Juhasz, N. Wallaszkovits, M. Födisch). Die „Re-Archivierung" dieses Teilbestands der Tonbandsammlung wird – soweit noch möglich – ausgehend von den Originalbändern, die der Urheber seinerzeit aus Spargründen als Arbeitskopien erhalten hatte, durchgeführt. Zu diesem Zwecke konnten 101 Originalbänder, die K. Gaál 1986 als Teil seines Vorlasses dem Komitat Vas zur Deponierung im Savaria Múzeum in Szombathely überantwortet hatte, nach zeitaufwendigen Verhandlungen mit demselben ausgeborgt werden. Im Zuge dessen wurden die nunmehrigen ungarischen Eigentümer des Bandmaterials von der Rechtsabteilung der ÖAW (Mag. Santoll) über die Rechte an diesen bereits im Phonogrammarchiv archivierten Aufnahmen, die gänzlich bei der ÖAW und dem Phonogrammarchiv liegen, schriftlich in Kenntnis gesetzt. Bis Ende des Jahres konnten die Originalbänder fast vollständig digitalisiert werden.

2.5. Übernahme und Aufarbeitung von Sammlungen

Der zweite Teil der Sammlung „Alfred Jilka, Vogelstimmen (1964–1983)" (31 Nagra-SN-Bänder, 15 Stunden) wurde digitalisiert (M. Födisch, N. Wallaszkovits) und mit der Archivierung sämtlicher Aufnahmen begonnen (K. Thenius-Wilscher).

Die Archivierung der bereits digitalisierten Sammlung „Herrmann Jungraithmayr" konnte aufgrund der Personalkürzungen (Einzeltransfer C. Gütls an die Universität Wien) nicht durchgeführt werden. Die weitere Aufarbeitung des Nachlasses von Prof. Anton Vorbichler musste ebenfalls hintangestellt werden.

Im Zuge des bis 31.03.2012 am Phonogrammarchiv laufenden Projekts „Das Phonogrammarchiv – Spiegel österreichischer Afrikaforschung" (gefördert von der Gem. Wien, MA 7) wurden dem Archiv zwei Tonbandsammlungen inklusive begleitender Bild- und Textmaterialien überantwortet: Die ersten Teile der „Sammlung Peter Fuchs" (Feldforschungsaufnahmen der Afrika-Expeditionen Tschad und Sudan 1955–1964, Niger 1972–1984, Tschad 1995/96) sowie der „Sammlung Andreas Kronenberg" (Aufnahmen aus Sudan 1957, 1958–1965 und in den 1980er-Jahren) wurden bereits übernommen (C. Gütl, H. Köb). Der Anteil „Sudan 1957" der „Sammlung Kronenberg" konnte digitalisiert werden (N. Wallaszkovits, M. Födisch).

2.6. Datenbank und Online-Katalog

In der Dokumentationsdatenbank wurde das Modul zur Verwaltung der Archivbenutzung (Nachweis der vergebenen Kopien, Benutzerdaten etc.) ausgetestet und in das Live-System integriert. Das Content Management System (CMS) wurde außerdem um Eingabefelder für Geodaten (Längen- und Breitengrad) erweitert, um im Online-Katalog abrufbare Bestände in der Soundmap des EU-Projekts „European Acoustic Heritage" zu integrieren (s. 2.2.). Die Ergänzung bzw. Verbesserung von Metadaten der historischen, bereits in der CD-Edition erschienenen Tonaufnahmen sowie die Verknüpfung der entsprechenden Wave-Files mit der Datenbank wurde in Angriff genommen (Volontariat Elisabeth Trojan). Mit dem Ziel, die Dokumentations- und Archivierungsarbeit zu ökonomisieren, wurde mit der Entwicklung eines Dokumentations-CMS für FeldforscherInnen begonnen, dessen Datensätze in die Archiv-Datenbank automatisch übernommen werden können, wodurch – idealiter – von den ArchivmitarbeiterInnen nur mehr präservatorische und technische Metadaten ergänzt werden müssen. Eine Beta-Version des „Feld-CMS" wird derzeit von mit dem Phonogrammarchiv kooperierenden Feldforscherinnen getestet, um basierend auf deren Feedback Verbesserungen vornehmen zu können. (C. Fennesz-Juhasz, M. Hagleitner)

3. Technische Forschung und Entwicklung im audiovisuellen Bereich

3.1. Konservatorisch-restauratorische Grundlagenforschung

Die international herausragende Stellung des Phonogrammarchivs im Bereich der technisch-konservatorischen Langzeitsicherung von audiovisuellen Datenträgern konnte durch die rezenten Forschungen im Berichtsjahr ausgebaut werden. Nach der 2011 erfolgten PCT Patentanmeldung mit dem Titel „Verfahren zur Rekonditionierung von Datenträgern", PCT/AT2011000516, einer Methode zur Rekonditionierung extrem gealterter und völlig unspielbarer Tonbänder auf Cellulose-Acetatbasis (Verfahren zur Rekonditionierung von Datenträgern Nr. A2124/2010, 23.12.2010), wurden von Nadja Wallaszkovits unter Einbezug von Fachleuten des Österreichischen Forschungsinstituts für Chemie und Technik (OFI) weitere Erwiderungen auf die Prüfungsbescheide des internationalen Patentamtes ausgearbeitet.

Die Methode kann derzeit aufgrund fehlender Mittel nicht in gewünschtem Maße weiterentwickelt werden. An der Drittmittelakquise in Kooperation mit einem nationalen Partner aus der Wirtschaft (Einreichung bei der FFG im Rahmen des BRIDGE Programmes) wurde im Projektjahr gearbeitet. Die Methode traf auch international bereits auf höchstes Interesse. Die

Zusammenarbeit mit einem der größten kommerziellen Archive weltweit wurde angebahnt.

Die Forschungen zur Extraktion des Vormagnetisierungssignals zur Korrektur von Gleichlaufschwankungen bei analogen Tonbandaufnahmen wurde insofern für Spezialfälle in den internen Archivworkflow als Prototyp eingebaut, als sie bei einzelnen Bändern aus dem EU-TEMPUS-Projekt „InMuSWB“ (s. 2.2.) versuchsweise angewendet werden konnte. Die praktische Implementierung in den täglichen Archivworkflow als Marktlösung für die internationale Archivgemeinschaft bräuchte aber noch weiterführende Forschungs- und Entwicklungstätigkeit und scheitert derzeit am Mangel finanzieller Mittel.

Die Arbeiten von Nadja Wallaszkovits an ihrer Dissertation zum Thema „Restaurierung historischer Audiomaterialien mit Schwerpunkt Feldforschungsaufnahmen aus dem Phonogrammarchiv der ÖAW“ konnten nicht ausreichend weitergeführt bzw. beendet werden – nicht zuletzt aufgrund der allgemein extrem hohen Auslastung der Audiotechnik, welche durch Personalkürzungen und Karenz derzeit nur mehr aus einer Person besteht.

Im Bereich der Feldforschungsberatung wurden experimentelle Vergleichsaufnahmen zwischen stereophoner Aufnahme und Mehrkanaltechnik weiter fortgeführt und entsprechendes Equipment bereitgestellt.

Im Rahmen des EU-Projekts „European Acoustic Heritage“ wurde von Jürgen Schöpf, Nadja Wallaszkovits und Bernhard Graf eine erste Vergleichsstudie verschiedener Aufnahmetechniken durchgeführt und die Ergebnisse in Hinblick auf praktische Anwendbarkeit ausgewertet (s.u., Publikationen).

3.2. Digitalisierungsprojekte, Re-Recording

Im Jänner 2012 erfolgte die konservatorische Begutachtung der Tonbandsammlung der Bayerischen Staatsbibliothek (Nadja Wallaszkovits). Es wurde eine umfassende Beratung, auch in Hinblick auf eine mögliche zukünftige Zusammenarbeit bei der Digitalisierung stark geschädigter Medien, durchgeführt.

Im März 2012 wurden 11 Tonbänder aus den Beständen des Thomas Bernhard Archivs in Zusammenarbeit mit MMag. Olivia Vrabl (Dissertantin an der Universität Wien) digitalisiert.

Weiters wurde die Sicherung und Digitalisierung der Tonbandsammlung des Oberösterreichischen Landesarchivs im Berichtsjahr zur Hälfte abgeschlossen. Es handelt sich dabei um 96 Tonbänder mit Aufnahmen von Landtagssitzungen, vorwiegend im Consumerformat Viertelspur und mit extrem langsamen Bandgeschwindigkeiten aufgenommen. Bei diesem Digitalisierungsauftrag zeigt sich die technische Vorrangstellung des

Phonogrammarchivs insofern, als diese Formate österreichweit nur mehr hier in höchster Qualität wiedergegeben und digitalisiert werden können.

3.3. Videographie, technische Sicherung von Videobeständen, Netzwerk

Nach den Vorarbeiten zur Einleitung der Migration wurde durch die Reduktion des Personalstandes und die Übersiedlung von Prof. Brandl eine vorübergehende Änderung der Aktivitäten erforderlich. Zeitkritischen Routinearbeiten musste der Vorrang eingeräumt werden. So wurde die umfangreiche Sammlung Brandl, die kurzfristig einen erheblichen Teil des Speicherplatzes des PhA-Servers belegte, von Franz Pavuza und Li Huang in eine Form gebracht, die eine Weiterverwertung erleichtert. Dazu mussten etwa 700 Stunden Videomaterial selektiert, einheitlich bezeichnet, von Mehrfachkopien befreit und auf externe Speicher der RMB-Collection (Arbeitsgruppe Prof. Rudolf M. Brandl) übertragen werden. Die ISO-Dateien der RMB-Collection sind vollständig gesichert und zugänglich, die zugehörigen geschnittenen AVI-Dateien werden derzeit von Prof. Brandl in der Bezeichnung vereinheitlicht und in ein übersichtliches System von Dateiordnern eingebracht, von wo sie auf Sicherungsbänder übertragen werden. Li Huang setzte ihre Arbeiten zur Katalogisierung der Originale fort (Bandmaterial; soweit bereits verfügbar).

Somit wurde seitens des Phonogrammarchivs für Ordnung, digitale Sicherung und infrastrukturelle Unterstützung der RMB-Collection eine Arbeitsleistung (insbesondere von F. Pavuza und M. Risnyovszky) im Umfang von ca. 200 Stunden erbracht.

Die für das Archiv anfallenden Routinearbeiten (Kopien, Extraktion von Bild und Ton; Video- und Bildbearbeitungen für Publikationen und Vorträge) wurden durch die Reduktion des Personals (Kündigung von Bernhard Graf) beeinträchtigt.

Die Sicherung des Videomaterials unterstützter FeldforscherInnen wurde kontinuierlich fortgesetzt, ebenso die Optimierung des Workflows zur Videoarchivierung (Hedwig Köb).

Im dritten Quartal führte Michael Risnyovszky einen groß angelegten Umbau des Netzwerks durch, um die Anbindung der Computer von Prof. Brandl und Li Huang am Fleischmarkt an das Phonogrammarchiv zu ermöglichen. In diesem Zusammenhang wurde auch die Konfiguration der Institutsfirewall überarbeitet und das Betriebssystem der Firewall auf den neuesten Stand gebracht. Im vierten Quartal musste das Netzwerk ein weiteres Mal umgebaut werden, diesmal, um die Anbindung des Phonogrammarchivs an das Akademie-Rechenzentrum vorzubereiten.

Publikationen

Selbstständige bzw. herausgegebene selbstständige Publikationen

Fennesz-Juhasz, Christiane, Petra Cech, Mozes F. Heinschink & Dieter W. Halwachs (Hg.). 2012. *Lang ist der Tag, kurz die Nacht: Märchen und Erzählungen der Kalderaš. / Baro o djes, cîni e rjat: Paramiča le Kaldêrašengê*. Klagenfurt: Drava.

Huang, Li (Red.). 2012. 有声遗产的维护: 规范、原则与保存策略. [Chinesische Übersetzung von: Dietrich Schüller et al. (eds.). *The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy.* (IASA Technical Committee – Standards, Recommended Practices and Strategies, IASA-TC 03). Version 3, December 2005. Online: <http://iasa-web.org/sites/default/files/downloads/publications/TC03-Chinese-simplified.pdf>]

Lechleitner, Gerda & Christian Liebl (Hg.). 2012. *Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften* 3. Göttingen: Cuvillier.

Beiträge in Fachzeitschriften und Sammelwerken

Cech, Petra, Christiane Fennesz-Juhasz & Mozes F. Heinschink. 2012. „Märchen und Erzählungen der Kalderaš". In: Fennesz-Juhasz, Christiane et al. (Hg.). *Lang ist der Tag, kurz die Nacht: Märchen und Erzählungen der Kalderaš. / Baro o djes, cîni e rjat: Paramiča le Kaldêrašengê*. Klagenfurt: Drava, 507–564.

Kowar, Helmut. 2012. „P. Primitivus Niemecz (1750–1806): seine ‚musikalischen Spieluhren und Maschinen'". *Das mechanische Musikinstrument* 115 (Dezember 2012): 7–13.

Kowar, Helmut. 2012. „Musik als Experiment? Zu Haydns Stücken ‚für das Laufwerk'". In: Dittrich, Marie-Agnes, Martin Eybl & Reinhard Kapp (Hg.). *Zyklus und Prozess. Joseph Haydn und die Zeit*. Wien: Böhlau, 293–304.

Kowar, Helmut. 2012. "Videotaping a Musical Box: An Attempt to Transcribe and Analyze Its Music". *Mechanical Music – Journal of the Musical Box Society International* 58 (1): 19–22.

Kowar, Helmut. 2012. "The trumpeter automaton of 1582: bringing its music back to life / Automatul cu trompetişti din 1582: Re(aducerea) în acutalitate a muzicii instrumentale". *Buletinul Muzeului Ştiinţei şi Tehnicii "Ştefan Procopiu" Iaşi* VI (6): 21–30.

Lechleitner, Gerda. 2012. "'Nidky nesbírame 'píseň', ale jen jednu její variantu' – úvahy o raných nahrávkách lidové hudby" ["'We never collect 'a song', but only one version of it' – Reflections on Early Folk Music Recordings"]. In: Procházková, Jarmila (ed.). *Vzaty do fonografu. Slovenské a moravské písně v nahrávkách Hynka Bíma, Leoše Janáčka a Františky Kyselkové z let 1909–1912* [*As Recorded by the Phonograph. Slovak and Moravian Songs Recorded by Hynek Bím, Leoš Janáček and Františka Kyselková in 1909–1912*]. Praha: Etnologický ústav AV ČR, v. v. i., 43–58.

Lechleitner, Gerda. 2012. "'Old wine in new bottles' – maintaining tradition through modern interpretation". In: Hemetek, Ursula (ed.). *Music and Minorities in Ethnomusicology: Challenges and Discourses from Three Continents*. Wien: Institut für Volksmusikforschung und Ethnomusikologie, 105–111. [Online: <http://www.mdw.ac.at/upload/MDWeb/ive/pdf/Music_an_Minorities_in_Ethnomusicology_-_Challenges_and_Discourses_from_Three_Continents.pdf>]

Lechleitner, Gerda. 2012. "The historical collections of the Phonogrammarchiv in Vienna – some reflections on dealing with presenting intangible heritage". In: Talam, Jasmina, Fatima Hadžić & Refik Hodžić (eds.). *7th International Symposium "Music in Society"*. Sarajevo: Musicological Society of the FB-H, Academy of Music in Sarajevo, 135–141.

Lechleitner, Gerda & Nona Lomidze. 2012. "Early sound documents of multipart music: concepts and historical context, analysis and interpretation". In: Macchiarella, Ignazio (ed.). *Multipart Music – a specific mode of musical thinking, expressive behavior and sound.* (Papers from the First Meeting of

the ICTM Study Group on Multipart Music, 15-20 September 2010, Cagliari, Sardinia). Udine: Nota, 37–50.

Lechleitner, Gerda & Nona Lomidze. 2012. "Some Considerations on the Interpretation of Multipart Music on Early Sound Documents". In: Tsurtsumia, Rusudan & Joseph Jordania (Hg.). *The Fifth International Symposium on Traditional Polyphony: Proceedings*. Tbilisi: International Research Center for Traditional Polyphony, Tbilisi State Conservatoire, 492–505.

Liebl, Christian. 2012. „Subak Julius (Giulio), Romanist und Lehrer". In: Österreichische Akademie der Wissenschaften (Hg.). *Österreichisches Biographisches Lexikon 1815–1950*. Wien: VÖAW, 21–22.

Pavuza, Franz. 2012. „Detailprobleme und Lösungen aus dem Digitalisierungsalltag eines kleinen Videoarchivs". In: Fröschl, Gabriele, Rainer Hubert, Elke Murlasits & Siegfried Steinlechner (Hg.). *Reale Probleme und virtuelle Lösungen. Eine Bestandsaufnahme anlässlich 50 Jahre Österreichische Mediathek und des UNESCO-World-Day for Audiovisual Heritage 2010. Gerhard Jagschitz zum 70. Geburtstag*. Münster/Wien: LIT-Verlag, 51–58.

Pavuza, Franz. 2012. „10 Jahre Videographie am Phonogrammarchiv". In: Lechleitner, Gerda & Christian Liebl (Hg.). *Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften* 3. Göttingen: Cuvillier, 26–34.

Schabus, Wilfried. 2012. „Ein Schatz an Wörtern und der Schatz im Wort. Impressionen aus Wortschatzerhebungen". In: Glauninger, Manfred Michael & Bettina Barabas (Hg.). *Wortschatz und Sprachkontakt im Kontext oberdeutscher Wörterbücher, Sprachatlanten und Sprachinseln: Werner Bauer zum 70. Geburtstag. Beiträge zur internationalen Tagung, Wien, 20. November 2009. / Ortsgrammatiken als Unterrichtsbehelf: „Laiengrammatiken" für Minderheitensprachen. Beiträge zum internationalen Workshop, Wien, 21. November 2009.* (Beiträge zur Sprachinselforschung, 21). Wien: praesens, 133–142.

Schöpf, Jürgen & Gabriele Proy. 2012. *Soundscape work in Austria – an overview*. Online: <http://europeanacousticheritage.eu/2012/09/soundscape-work-in-austria-an-overview/>.

Schöpf, Jürgen, Christiane Fennesz-Juhasz & Dietrich Schüller. 2012. "Acoustic Heritage in the Phonogrammarchiv of the Austrian Academy of Sciences". In: Kytö, Meri, Nicolas Remy & Heikki Uimonen (eds.). *European Acoustic Heritage*. Tampere: Tampere University of Applied Sciences, 42–52. [Online: <http://europeanacousticheritage.eu/wp-content/uploads/2012/09/EuropeanAcousticHeritage_web.pdf>]

Schöpf, Jürgen, Christian Liebl & Christiane Fennesz-Juhasz. 2012. "The Phonogrammarchiv at the Austrian Academy of Sciences". In: Kytö, Meri, Nicolas Remy & Heikki Uimonen (eds.). *European Acoustic Heritage*. Tampere: Tampere University of Applied Sciences (TAMK), 20–24. [Online: <http://europeanacousticheritage.eu/wp-content/uploads/2012/09/EuropeanAcousticHeritage_web.pdf>]

Schöpf, Jürgen, Nadja Wallaszkovits & Bernhard Graf. 2012. "Practicality of location recording in surround technology – exploring a workflow from field recording to user-friendliness to long-term preservation". In: Kytö, Meri, Nicolas Remy & Heikki Uimonen (eds.). *European Acoustic Heritage*. Tampere: Tampere University of Applied Sciences (TAMK), 86–95. [Online: <http://europeanacousticheritage.eu/wp-content/uploads/2012/09/EuropeanAcousticHeritage_web.pdf>]

Spitzbart, Johannes. 2012. „Phonogrammarchiv goes DISMARC. Hintergründe zur Migration des Online-Kataloges des Phonogrammarchivs". In: Lechleitner, Gerda & Christian Liebl (Hg.). *Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften* 3. Göttingen: Cuvillier, 22–25.

Wallaszkovits, Nadja. 2012. „Konditionierung gealterter Magnetbänder auf Cellulose-Acetatbasis". In: Fröschl, Gabriele, Rainer Hubert, Elke Murlasits & Siegfried Steinlechner (Hg.). *Reale Probleme und virtuelle Lösungen. Eine Bestandsaufnahme anlässlich 50 Jahre Österreichische Mediathek und des UNESCO-World-Day for Audiovisual Heritage 2010. Gerhard Jagschitz zum 70. Geburtstag*. Münster/Wien: LIT-Verlag, 43–50.

Wallaszkovits, Nadja. 2012. „Vom Informationsträger zum Kunstwerk – Ein Blick über den (Band-) Tellerrand“. In: Murlasits, Elke & Gunther Reisinger (Hg.). *Museum multimedial. Audiovisionäre Traditionen in aktuellen Kontexten*. Münster/Wien: LIT-Verlag, 171–182.

Wallaszkovits, Nadja. 2012. „Zur Digitalisierung der Sammlung Armando Leça“. In: Lechleitner, Gerda & Christian Liebl (Hg.). *Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften* 3. Göttingen: Cuvillier, 9–21.

In Kooperation mit dem Phonogrammarchiv entstandene CD-Publikationen

Aschauer, Mario. 2012. *Schubert Adrast – Einzelnummern und Entwürfe zu einer Oper D 137*. (Klingende Forschung 1, Hg.: Kommission für Musikforschung; OEAW PHA CD 30). Wien: VÖAW. [J. Spitzbart & N. Wallaszkovits: Aufnahme; N. Wallaszkovits: Mastering; G. Lechleitner: Redaktion]

Klugseder, Robert & Franz Karl Praßl. 2012. *Gregorianischer Choral aus Quellen österreichischer Bibliotheken.* (Klingende Forschung 2, Hg.: Kommission für Musikforschung; OEAW PHA CD 31). Wien: VÖAW. [J. Spitzbart: Aufnahme, Mastering; G. Lechleitner: Redaktion]

Rausch, Alexander. 2012. *Romanorum rex. Musik in Zentraleuropa um die Mitte des 15. Jahrhunderts.* (Klingende Forschung 3, Hg.: Kommission für Musikforschung; OEAW PHA CD 32). Wien: VÖAW. [J. Spitzbart & N. Wallaszkovits: Aufnahme; N. Wallaszkovits: Mastering; G. Lechleitner: Redaktion]

Veröffentlichungen und Feldforschungsaufnahmen der RMB-Collection (Orbis Musicarum)

Brandl, Rudolf M. & Bernhard Graf. 2012. „Griechenland-Feldforschung 2011 im Epiros und auf Karpathos”. In: Lechleitner, Gerda & Christian Liebl (Hg.). *Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften 3*. Göttingen: Cuvillier, 81–93.

Brandl, Rudolf M. 2012. Panegyri in Nikopolis Prevezas 2011, 3 DVD (= OM 165); Panegyri mit Trapeza in Koukli/Pogoni 2011, 3 DVD (= OM 166); Panegyri Agias Marinas (Olympos/Karpathos) 2011, 2 DVD (= OM 167); Festaufführung (6 Einakter) der Kunqu-Liebhaber am Kunqu-Festival 2011 in Beijing (= OM 168); „Qishuanhui“– Aufführung der Beifang Kunqu Juyuan beim Kunqu-Festival 2011 (= OM 170); „Hongloumeng“ 2. Teil (Traum der Roten Kammer), moderne Kunqu-Oper, Aufführung der Beifang Kunqu Juyuan beim Kunqu-Festival 2011, 2 DVD (= OM 171); Sarakatsanische Hochzeit in Kalpaki/Epiros 2011, 3 DVD (= OM 172); Musikerhochzeit der Roma in Ioannina 2011, 3 DVD (= OM 173); Gastspiel der Beifang Kunqu Juyuan in Wien 2011: Lankeshan (= OM 174); Gastspiel der Beifang Kunqu Juyuan in Wien 2011: Dou E Yuan (= OM 175); Kanton-Oper (Yueju): „Zuida Jinzhi“ (Betrunken die Gattin schlagen) im Nanfang-Theater Guangzhou 2012, 2 DVD (= OM 176); Hunan-Folklore-Aufführung 2012 in Linwu, Hunan für die Nuo Research Society (= OM 177); Österreichisches [*recte* Wiener] Volksliedwerk: „Des is ka Schmäh“– Kurt Girk zum 80. Geburtstag, Bockkeller (Wien) 2012 (= OM 178); Nuo-Maskenritual in Dachong Xiang, You Wancun, Kreis Linwu, Hunan. Sonderaufführung für die Nuo Research Society, 3 DVD (= OM 179); Nuoxi Exorzismus-Theater am Ende des Nuo-Maskenrituals in Dachong Xiang, You Wancun, Kreis Linwu, Hunan. Sonderaufführung für die Nuo Research Society (= OM 180); Säkulares Panegyri in Charokopi (Tzoumerka) 2011, 2 DVD (= OM 182); Festakademie der Vereinigung „Das Wienerlied“ zum 50. Geburtstag von Herbert Schönfelder 2003 (= OM 184); „Jetzt gemma’s an!“ – Thomas Hojsa & Helmut Emersberger bei der Wienerlied-Vereinigung „Robert Posch“ 2010 (= OM 185); „Wer net singt, der hot ka Herz im Leib“– Abend im Weinlokal „Zum G’spritzten“ 2012 in Ottakring mit dem Kurt Girk Trio (Kurt Girk, Rudi Koschelu, Thomas Hojsa) (= OM 186); „Zauber der Vorstadt“ – Abend im Hernalser Kultur-Café „Max“ 2012 mit Kurt Girk, Thomas Hojsa und Christoph Lechner (= OM 187); Traditional Bosnian Music and Sevdalinka in Sarajevo 2012: Ansambl Etnoakademik (= OM 188); Bosnian Tamburica Music live with Basstamburaši Bend in taverna „Mala Pivniva“ 2012 (= OM 189); Concert with Traditional Bosnian Folkmusic 2012 in Sarajevo (= OM 190; DVD, CD);

Chaldäisches Tanzfest am Ostersonntag 2009 in Wien, 2 DVD (= OM 191); Thede Kahl: Corinzi – Colinde. Weihnachtslieder und winterliche Heischebräuche aus dem Motzenland (Siebenbürgen) in Wien 2012 (= OM 192); Lili Hsueh (Taiwan): Gezaixi (Taiwanesische Lokaloper): „Shiwangji" (Der König der Löwen) 2004 im Bodhi-Kloster in Taiwan, 2 DVD (= OM 193). [Alle:] Göttingen: Cuvillier.

Vorträge

Fennesz-Juhasz, Christiane (10.10.2012): Curating for manifold use in a digital world (read by N. Wallaszkovits). IASA Annual Conference 2012, New Delhi.

Huang, Li & Rudolf Brandl (16.08.2012): Nuo in Anhui [in Chinesisch]. Internationale Konferenz der Nuo-Kultur, Linwu, Provinz Hunan/VR China.

Huber, Christian (24.04.2012): Subject and object agreement in Shumcho. Vienna discussion group in linguistics (Institut für Sprachwissenschaft, Universität Wien), Wien.

Huber, Christian (25.04.2012): Agreement marking in complement clauses in Shumcho. Vienna discussion group in linguistics (Institut für Sprachwissenschaft, Universität Wien), Wien.

Huber, Christian (10.09.2012): Transitivity alternations and causativisation in Shumcho. 18th Himalayan Languages Symposium, Varanasi/Indien.

Kowar, Helmut (07.09.2012): The trumpeter automaton of 1582: bringing its music back to life. Festivalul International al muzicii mecanice, Iaşi.

Kowar, Helmut (12.09.2012): Historische Musikautomaten – ein Fenster zur klingenden Musik der Vergangenheit. Sonderausstellung „Musik aus der Dose", Innsbruck.

Kowar, Helmut (15.09.2012): P. Primitivus Niemecz (1750–1806): seine „musikalischen Spieluhren und Maschinen". Jahrestagung der Gesellschaft für selbstspielende Musikinstrumente, Wien.

Lechleitner, Gerda (09.08.2012): The methodology justifies the methods. 7th Study Group Meeting Music and Minorities (Zefat Academic College), Zefat/Israel.

Lechleitner, Gerda (20.09.2012): Soldier songs of the Austro-Hungarian Army – "real" memory or "imagined" myth? European Seminar in Ethnomusicology XXVIII: Music and Cultural memory in post-1989 Europe (Center for Interdisciplinary Research and Institute of Ethnomusicology ZRC SAZU), Ljubljana.

Lechleitner, Gerda (28.11.2012): Die frühe Sammlung „fremder" Musikinstrumente durch die Gesellschaft der Musikfreunde. Symposium: Die Emporbringung der Musik. 200 Jahre Gesellschaft der Musikfreunde in Wien (Gesellschaft der Musikfreunde in Wien), Wien.

Lechleitner, Gerda, Franz Lechleitner & Nona Lomidze (28.09.2012): CD-Project: recordings from the Caucasian Region, 1909 and 1915/16. 6th International Symposium on Traditional Polyphony (International Research Center for Traditional Polyphony of Tbilisi State Conservatoire), Tbilisi.

Lechleitner, Gerda & Nona Lomidze (25.04.2012): Production and process – creative acting and perceiving. 2nd Symposium of the ICTM Study Group for Multipart Music, Tirana.

Liebl, Christian (24.02.2012): Brandl, Doegen and the recording campaign in German prisoner-of-war camps during World War I – a neglected resource for English dialectology. Symposium: Studying Change in the History of English – New Directions in Linguistics (Universität Wien; Birmingham City University), Wien.

Liebl, Christian (27.04.2012): 113 Jahre – und kein bisschen leise: Das Phonogrammarchiv und die Gesamtausgabe der Historischen Bestände. Lange Nacht der Forschung (Österreichische Akademie der Wissenschaften), Wien.

Remmer, Ulla (11.05.2012): Historical Sound Recordings of Brahman Sanskrit Recitation in the Phonogrammarchiv of the Austrian Academy of Sciences. The Collection of Felix Exner (1904–1905). Symposium: Iranische und indoeuropäische Onomastik und Sprachwissenschaft. In memoriam Manfred Mayrhofer, 1926–2011 (Institut für Iranistik, ÖAW), Wien.

Schöpf, Jürgen (10.05.2012): "Religion Road": a Series of Religious Sites in Changlang district, Arunachal Pradesh, India. ISA Guest Lecture (Institut für Sozialanthropologie, ÖAW), Wien.

Schöpf, Jürgen (06.09.2012): What's so special about music? – Some methodological aspects of ethnomusicological field data in the Phonogrammarchiv of the Austrian Academy of Sciences, Wien. Tagung: Best practices in digital language archiving of language and music data (Universität zu Köln, Institut für Linguistik), Köln.

Schöpf, Jürgen (19.09.2012): Phonogrammarchiv at the Austrian Academy of Sciences. Meeting: Discovering the Soundscapes of Europe – EAH Exhibition and publication seminar (Fachhochschule Tampere, Abteilung Kunst- und Mediengestaltung), Tampere.

Schöpf, Jürgen (20.09.2012): Recording Soundscapes. European Acoustic Heritage Conference (Fachhochschule Tampere, Abteilung Kunst- und Mediengestaltung), Tampere.

Schöpf, Jürgen (20.09.2012): Sound archives and re-studying soundscapes. European Acoustic Heritage Conference (Fachhochschule Tampere, Abteilung Kunst- und Mediengestaltung), Tampere.

Schüller, Dietrich (23.02.2012): Introduction to IASA-TC 05 Handling and Storage of Audio and Video Carriers. 2012 TELDAP International Conference, Taipeh.

Schüller, Dietrich (16.03.2012): Safeguarding the Audiovisual Heritage in SEE: The situation of Cultural and Research Collections. Vorbereitungstreffen für den Annual Council of Ministers of Culture in South East Europe, Belgrade.

Schüller, Dietrich (20.04.2012): Safeguarding the Audiovisual Heritage in South East Europe. Annual Council of Ministers of Culture in South East Europe, Belgrade.

Schüller, Dietrich (19.05.2012): Magnetic Tape Stability: Interviews with Representatives of Former Tape Manufacturers. ARSC 46th Annual Conference, Rochester, New York.

Schüller, Dietrich (19.06.2012): Langzeitsicherung audiovisueller Bestände: Prinzipien, Strategien, konkrete Planung. Nestor-Praktikertag 2012, Berlin.

Schüller, Dietrich (16.07.2012): UNESCO's Information for All Programme (IFAP), Multilingualism and Cultural Diversity. International Workshop UNESCO-Chair in Multilingual Transcultural Communication in the Digital Age, Universität Wien, Wien.

Schüller, Dietrich (14.09.2012): Standards und Strategien zur Langzeitarchivierung audiovisueller Datenträger. ODOK 2012, Wels.

Schüller, Dietrich (28.09.2012): Challenges for the Preservation of Audiovisual Documents. UNESCO-Tagung "The Memory of the World in the Digital Age", Vancouver.

Schüller, Dietrich (07.10.2012): Talking to Former Tape Manufacturers. IASA Annual Conference 2012, New Delhi.

Wallaszkovits, Nadja (27.09.2012): Digitisation of Small Sound Collections: Problems and Solutions. UNESCO Tagung "The Memory of the World in the Digital Age", Vancouver.

Wallaszkovits, Nadja (10.10.2012): Providing Access to Research Sound Collections – "Quick and Dirty" or Highest Quality? IASA Annual Conference 2012, New Delhi.

Wallaszkovits, Nadja (29.10.2012): The Replay of Historical Magnetic Tape – More than Pressing the Play Button. 133rd AES Convention 2012, San Francisco.

Wallaszkovits, Nadja & Friedrich Engel (24.11.2012): Die Digitalisierung der Sammlung Armando Leça: Ein Beitrag zur Geschichte der Magnetbandaufnahme im Feld 1939–40. 27. Tonmeistertagung, VDT International Convention, Köln.

Wallaszkovits, Nadja & Heinrich Pichler (24.11.2012): Probleme mit Rauschunterdrückungsverfahren bei der Archivierung analoger Consumer-Tonbandformate. 27. Tonmeistertagung, VDT International Convention, Köln.

Wallaszkovits, Nadja & Dietrich Schüller (26.04.2012): The Replay of Historical Magnetic Tape – More than Pressing the Play Button. 132nd AES Convention 2012, Budapest.

AUTOR/INN/EN

Vanna Viola Crupi received her PhD in history and analysis of musical cultures from the Sapienza University of Rome in 2011. Since 2005 she has been a member of the Italian Ethnological Mission in Equatorial Africa.

Rinko Fujita received her PhD from the University of Vienna in 2007. Since 2010 she has been a lecturer in ethnomusicology at the University of Vienna.

Catherine Ingram is currently a Newton International Fellow at the School of Oriental and African Studies (SOAS) at the University of London, and has been an Honorary Fellow at the Melbourne Conservatorium of Music, University of Melbourne, Australia.

Gisa Jähnichen is Professor of Comparative Musicology and Audiovisual Media at the Music Department, Faculty of Human Ecology, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia.

Anastasia Karlsson received her PhD from the Centre for Languages and Literature, Lund University, Sweden. She has studied the underlying differences between prosodic systems of tonal and non-tonal languages on the basis of Kammu speech data.

Mag.a Barbara Kazianka studierte in Wien Sozial- und Kulturanthropologie und ist freischaffende Lektorin. Ihre Feldforschungen in den Jahren 2005/6, 2008, 2009, 2012/13 fanden im Rahmen der Dissertation zum Thema „Die Rolle von traditioneller Medizin in der Konstruktion indigener Identitäten anhand der Itzá Maya in Guatemala und der Wayuu in Kolumbien“ statt.

Bob Ladd is Professor Emeritus of Linguistics in Edinburgh, UK. An important part of his professional life is devoted to editorial work, e.g. for *Language and Speech* (1994–2000), *Phonology* (since 1995), *Linguistics*, *Journal of Linguistics* and *Journal of Phonetics*.

David Locke is Professor in the Music Department at Tufts University, USA. His doctoral dissertation is about the Ewe masterwork *Atsiagbekor* (Wesleyan University 1979). He is the author of three monographs and the African chapter in the textbook *Worlds of Music*.

Håkan Lundström is Professor at the Malmö Academy of Music, Lund University, Sweden. He has conducted research in Japanese and Alaskan Native American musics, but specialises in Kammu music, particularly the singing traditions.

Stephen Morey is Australian Research Council Future Fellow at the Centre for Research on Linguistic Diversity (CRLD) at La Trobe University, Melbourne. He has worked on endangered languages of North East India since 1999.

Murray Schellenberg received his doctorate in Linguistics from the University of British Columbia with a thesis on "The Realization of Tone in Singing in Cantonese and Mandarin". He is currently a post-doctoral associate at the University of Rochester, NY, USA.

Gretel Schwörer-Kohl has studied musicology, anthropology and history of East Asian art at the universities of Freiburg, Cologne and Hamburg. Since 1999 she has been Professor of Ethnomusicology at the Martin-Luther-Universität Halle-Wittenberg, Germany.

Jan-Olof Svantesson obtained his PhD in General Linguistics in 1983. He is Professor emeritus at the Centre for Languages and Literature, Lund University, Sweden, and an expert on Yùan Kammu and long-term coordinator of linguistic research on Kammu.

Siri Tuttle is Associate Professor of Linguistics at the University of Alaska Fairbanks, USA. Her dissertation was on the Tanana Athabascan language of interior Alaska, and she continues to study Alaskan Athabascan prosody.

Amanda Villepastour is a graduate of the University of Western Australia (BMus, composition) and the School of Oriental and African Studies, University of London (MMus, PhD ethnomusicology), where she completed her PhD in 2006. She is currently Lecturer in Ethnomusicology at Cardiff University.

GASTHERAUSGEBER

Jürgen Schöpf: Ethnomusikologe, Phonogrammarchiv der ÖAW

HERAUSGEBER/IN

Gerda Lechleitner: Musikwissenschaftlerin, Phonogrammarchiv der ÖAW
Christian Liebl: Anglist, Phonogrammarchiv der ÖAW