Discrete level narrative, terraced music: insights from underdocumented Ivorian languages

Dafydd Gibbon, Firmin Ahoua,
Blé François Kipré & Sascha Griffiths

Proceedings of Conference on

Language Documentation & Linguistic Theory 2

Edited by Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells

13-14 November 2009 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
http://www.soas.ac.uk/academics/departments/linguistics

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4578
Fax. +44-20-7898-4349
elap@soas.ac.uk
http://www.hrelp.org

© 2009 Dafydd Gibbon, Firmin Ahoua, Blé François Kipré & Sascha Griffiths

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0392-9

This publication can be cited as:

Dafydd Gibbon, Firmin Ahoua, Blé François Kipré & Sascha Griffiths. 2009. Discrete level narrative, terraced music: insights from underdocumented Ivorian languages. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS.

or:

Dafydd Gibbon, Firmin Ahoua, Blé François Kipré & Sascha Griffiths. 2009. Discrete level narrative, terraced music: insights from underdocumented Ivorian languages. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS. www.hrelp.org/eprints/ldlt2_11.pdf

Discrete level narrative, terraced music: insights from underdocumented Ivorian languages

DAFYDD GIBBON[1], FIRMIN AHOUA[2], BLÉ FRANÇOIS KIPRÉ[2] & SASCHA GRIFFITHS[3]
Universität Bielefeld[1], Université de Cocody, Abidjan, Côte d'Ivoire[2], University of Canterbury[3]

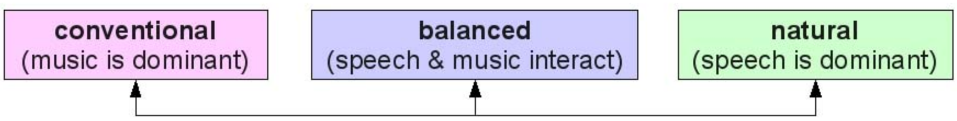
1. MELODY IN SPEECH AND SONG

This study addresses the similarities and differences between pitch contours in speech and the melody of song, based on documentation and analysis of data instances from two Niger-Congo (Atlantic-Congo) languages, the Daloa variety of Bete (Kru, ISO 693 *bev*) and the Sanvi variety of Anyi (Kwa, ISO 639-3 *any*) in a cooperation between three linguists, two of whom are experienced amateur musicians, and a musicologist who is a professional musician.

The relations between properties of speech and properties of music are complex, and only a small segment can be dealt with here. In particular, relations between speech music and grammar are excluded, and of the two main dimensions of music, rhythm and melody, only melody is treated. The melodic subdimension of intonation in the wider sense is also excluded, though it is also important for tone languages (Ahoua 1998) and a major component of studies in prosodic typology (Hirst & al. 1998). Further, there are clearly many styles of speech and genres of music, so results from a specific study cannot be easily generalised, but can only have the status of a reference model for further studies.

The main hypothesis of this study is that there is a Conventional-Natural Scale (Figure 1) for each of the dimensions of rhythm and melody, which together define a Conventional-Natural Space within which both speech styles (note the difference between formal and spontaneous speech) and musical genres (cf. baroque song vs. much popular music) are located.

Figure 1
Conventional - natural scale of speech-music relationships



The data are taken from two Niger-Congo tone languages with different typological prosodic features. Bete, a 4-tone language, is taken as a possible benchmark for discrete-level tone languages, a property associated with the often encountered conjecture which states (in simplified form) that register tone languages with two tones are susceptible to terracing (automatic downstep tone sequencing), while languages with more than two tones are much less susceptible to terracing or even downdrift of a more universal kind, and have tones which, roughly speaking, have their own non-overlapping or marginally overlapping spaces within the pitch range. This claim has not been verified in the literature on Bete, consequently a brief study of Bete tone patterning is included. Anyi is a 2-tone language, with pronounced tone-terracing. For detailed discussion of downtrend types cf. Connell (2002). For comparison purposes, Bete acts as a benchmark for comparison with possible song types with discrete pitch levels.

The main focus of the study is on the Awosi (dirge) variety of unaccompanied solo song in Anyi, and the quantitative relations between the pitch contour of the song and the pitch contours of male and female spoken versions of the song text.

The null hypothesis is obviously that speech and song melodies are identical. It is trivial. It is equally obvious that this hypothesis can be trivially refuted. However, this would be an uninteresting result. The interesting task is to find out precisely where there are differences and similarities between speech and song melodies, and whether these differences and similarities can be systematically accounted for in a parametric model which, beyond lingustic requirements, is of sufficiently high quality to be operationalised in a speech synthesis system.

In the literature, different authors, looking at different song genres and languages, have come to different conclusions about similarities and differences between speech and song (Agawa 1998; Aka 2007; Blacking 1973; Connell 2005; Greenberg 1949; Leben 1981; Leben 1985; Richard 1972; Schneider 1943-1944; Schneider 1961; Simmons 1980). However, the phonetic methodology developed for the present study for exact analysis has so far not been deployed elsewhere.

In the following sections, first the tone systems of Anyi (the Sanvi/Sanwi variey) and Bete (the Daloa variety) are briefly outlined, and the results of the analysis are presented. Finally, in the concluding section results and some consequences for future work are summarised.

2. TONE INVENTORIES OF BETE AND ANYI

The tonal inventories of Bete and Anyi are shown in Table 1 and Table 2, respectively. Further discussion of tones in context is not necessary in this context but can be obtained from Kipré (1989) and Marchese (1983) for Bete and Hérault (1982) for Anyi.

Table 1
The Bete tone inventory

Tone	**Example**	**Gloss**
Extra high	sű	‘dance’
High	sú	‘tree’
Mid	sū	‘sting’ (insect)
Low	sù	‘transport’

Table 2
Anyi tones in sequence

[fɔ̀ʊ̰̀]	(L L)	touchy	[fɔ́ʊ̰́]	(H H)	right	
[àȋɛ]	(L L H)	food	[àlɪ́ɛ̀]	(L H L)	day	
[bɔ̀lɛ]	(L L)	potash/ venom	[bɔ̀lɛ́]	(L H)	fog	[bɔ́lɛ̀] (H L) packet
[ɛ̀sɛ́]	(L LH)	then	[ɛ̀sɛ̂]	(L HL)	pot	
[èbô]	(L HL)	river	[èbǒ]	(L LH)	field	

3. METHOD

The methodology follows the relevant recommendations in Gibbon & al. (1997) and Gippert & al. (1996) and is related to the computational modelling techniques introduced in Gibbon & al. (2003).

For comparing speech and song at the phonetic level, the pitch contours of the speech and song instances in the data were extracted by taking the mean fundamental frequency (F0) of the vowel in the tone-bearing unit (syllable), in order to define a relation between underlying lexical tones and pitch contours. Only lexical register tones, not lexicalised contours or phonetic allotonic contours, were considered. The use of mean F0 is justified by the analysis of the tones as register tones; clearly a more detailed analysis of tones in context with a more highly structured phonetic representation will be needed in future work. For the comparison itself, it was considered sufficient to provide basic descriptive and linear regression statistics for the pitch contours. Again, more complex modelling criteria could be used, but were not necessary for the purpose at hand. For reasons of space, the actual texts recorded are not given.

The phonetic method for prosodic analysis was as follows.

(1) Data documentation: a selection of data for the study of pitch contours in spoken Bete and in spoken and sung Anyi was designed, recorded under relatively natural conditions, and documented together with manually annotated utterance, phrase, morph, phone, tone and pitch tiers, using the Praat phonetic toolbox. The data items are: (i) Bete: a story told in formal style (one data instance)[1]. (ii) Anyi: unaccompanied song (Awosi dirge), with the two textually almost identical parts analysed separately, and female and male spoken versions of the song text (four data instances in total).

[1] We gratefully acknowledge the contribution of F. Kipré Blé for cooperation in producing and analysing the Bete data.

(2) The Praat data format ('TextGrid') was converted into a standard database format (CSV, character separated values) using an online software tool.[2]
(3) The format was further transformed into a more compact CSV format, from which sequences of pairs of time-stamps and pitch values were extracted separately for all H tones and L tones, as well as for the entire tone sequence.
(4) The pitch values were used to calculate basic descriptive statistics for each of the sequences.
(5) The pairs of time-stamps and pitch values were used to calculate the linear regression (intersect and slope) properties of the H, L and combined pitch contours for each data item.

An automatic speech synthesis method was used for quality control of the annotated data (Gibbon & Bachan, 2008); synthesis examples are available for demonstration but will not be discussed further at this point.

4. RESULTS

The results for Bete and Anyi are presented separately, because the roles of the two languages in the present approach are different, Bete as an example of a putative discrete level system, and Anyi as an example of a terraced system. For comparison purposes, Bete would serve as a reference point for discrete level music, Anyi for terraced (or at least strongly downtrending) music.

The tale was read aloud, recorded and provided manually with a time-aligned annotation, and tones in sequences of extra-high (HH), high (H), mid (M) and low (L) tones were paired separately with their pitch values. The resulting contours are showin in Figure 2.

4.1. Results: Bete

For investigating Bete as a possible benchmark case of discrete level tones, a tale was recorded in a formal style of reading aloud by an experienced educated speaker in order to reduce the effect of rhetorical influences on pitch range and height in spontaneous speech, and thereby to permit investigation of the role of lexical tones in longer sequences. The story chosen was the classic 'The North Wind and the Sun', one of Aesop's fables, used by the IPA as a standard data source. The choice is deliberately not a traditional Bete tale, which requires a complex configuration of narrator, responder and audience, and involves too many intonation-influencing variables to be of much use in the study of lexical tone sequencing.

The results of applying basic descriptive statistics and linear regression to these F0 (fundamental frequency) value sequences are shown in Table 3.

[2] http://wwwhomes.uni-bielefeld.de/gibbon/Forms/Python/PHONETICS/textgrid2csv.html

Table 3
Measurements of F0 trajectory properties in the Bete narrative

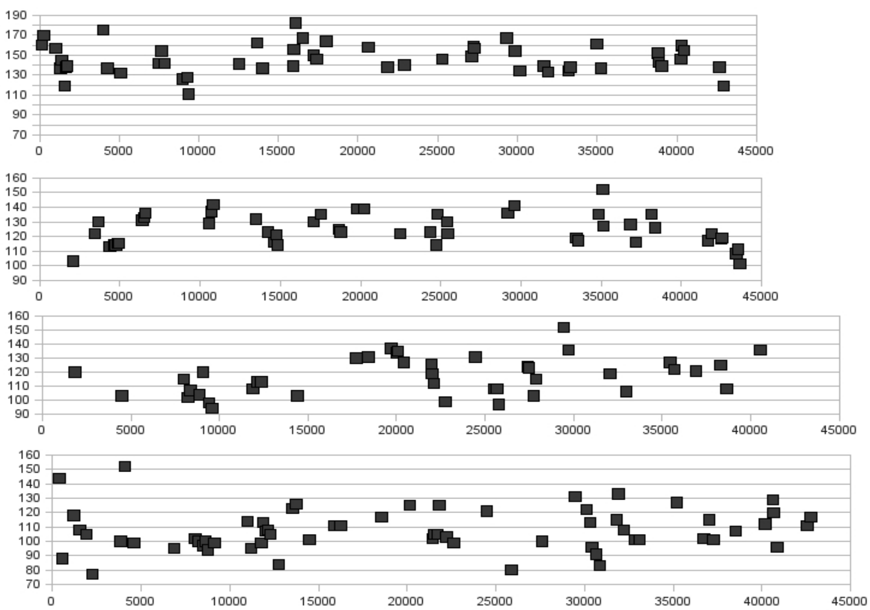
Tone	N	min	max	mean	sd	offset	slope
HH:	51	111	182	146	14	146	1.5e-05
H:	49	101	152	125	11	127	-5e-05
M:	41	94	152	117	13	107	0.0005
L:	62	77	152	107	14	105	0.0001

The calculations show that the means are clearly different and the standard deviations are small with little overlap. The intercepts and the means in each case are predictably similar. However, the slope in each case is indistinguishable from zero, indicating that at least at this global textual level each tone has its own discrete level, thus confirming the discrete level hypothesis for Bete.

4.2. Results: Anyi

Examination of the song texts shows that considerable elision of syllables (both L and H tone) takes place in comparison with the spoken versions, with a pronounced rhetorical effect. Likewise, timing is very different, for example with an overlong phrase-final [o:] appellative discourse particle in the song. An exact comparison therefore needs to take place at a stylised level which generalises over these differences.

Figure 2
Pitch contours for first and second Anyi song parts (top two contours, respectively), female and male speech (next to bottom and bottom, respectively)



The pitch contours for both sung and spoken conditions in Anyi are shown in Figure 3. The contours are not based on continuous pitch samples, but as already noted they are based on the average pitch on tone-bearing units; the points are stylised by means of a cubic spline function. The contours have not been completely normalised for syllable count; syllable elisions vary between data items.

Using an impressionistic 'eyeball method' to start with, male and female contours are seen to be conspicuously similar, and almost completely parallel. As expected, they are approximately 40 Hz apart. The song consists of two textually almost identical parts, and the first part is strikingly similar to the spoken versions. The same applies to the first portion of the second part of the song, but the second portion of this part is considerably modified, and less 'natural' than the first.

Confirmation of the eyeball impression is provided by the basic descriptive statistics and the regression analysis for Anyi speech and song, shown in Table 4 and Table 5, respectively.

Table 4
Pitch analysis for spoken versions of Anyi song

Speaker	Tone	N	Min	Max	Mean	sd	Intercept	Slope
Narcisse	H	7	91	142	115	17	146	-0.02
Narcisse	L	11	96	127	108	9.5	115	-0.006
Narcisse	H&L	18	91	142	111	14	130	-0.01
Rose	H	6	135	190	155	21	186	-0.02
Rose	L	11	127	185	150	16	160	-0.006
Rose	H&L	17	127	190	152	18	171	-0.01

Inspection of the results yields the following generalisations:

(1) The pitch contour of the female speaker is appreciably above (about 40 Hz on average) that of the male speaker, as shown by the mean and intercept values. This was to be expected.
(2) The slope values for both tones and the overall contour are identical, a somewhat surprising 'ideal' result. In a sense this was to be expected, but not to the extent which was actually found.
(3) Standard deviations show considerable phonetic overlap between H and L tones; further analysis is required in order to show whether the H and L contours are significantly different.

More detailed analysis of tone pairs in sequence, as well as a more structured model of pitch patterns on and between tone-bearing units are needed, but the present straightforward methodology is sufficient to demonstrate the consistency of patterning in the spoken versions of the song text.

Table 5
Pitch analysis for Anyi song (first part in the first three rows, second part in the last three rows)

Singer	Tone	N	Min	Max	Mean	sd	Intercept	Slope
Anon.	**H**	6	197	320	243	47	313	-0.035
Anon.	**L**	11	174	286	229	35	287	-0.034
Anon.	**H&L**	17	174	230	234	40	293	-0.33
Anon.	**H**	6	168	282	212	42	268	-0.029
Anon.	**L**	12	154	227	201	22	234	-017
Anon.	**H&L**	18	154	282	204	31	234	-0.017

The results for both parts of the song show both considerable differences and considerable similarities. The main difference is in pitch height: although the singer is male, F0 is higher throughout, with no overlap, even than the pitch of the female speaker. A spoken version by the singer is not available, but the difference is so great that a similar difference in song pitch and speech pitch would be expected.

Interestingly, for the present study, the overall slope of the pitch contour of the song is much greater than the slope of the pitch contour in the spoken data. Thus, the initial null hypothesis that the melodies of speech and song are the same is falsified *sensu strictu*, but *not* in the sense that the melody attains a form which is comparable to a discrete level system, since, on the contrary, the song adopts the downtrend feature of speech and amplifies this for rhetorical effect.

5. CONCLUSION

Expectations were fulfilled with regard to benchmark data for discrete levels, i.e. the four register tones and the discrete level hypothesis for Bete. It must be said, however, that the discrete level hypothesis was confirmed at the global textual rank, not at the rank of phrasal domains for which it is usually postulated. More work is needed before the model can be verified for smaller domains.

Expectations were also fulfilled in respect of the downtrend in Anyi speech. The 'eyeball' based intuition that speech and song melody are closely related in the case of the Awosi dirge was confirmed on the basis of regression analyses, of which the regression analyses were reported here. The Anyi song under consideration here is therefore located well towards the 'natural' end of the Conventional-Natural Scale in speech-song relationships. A caveat is necessary, however: first, linear regression is a highly simplified model, and, on the basis of what is known about pitch patterning, a logarithmic regression model would be required in order to get more exact results; second, detailed local analysis of tone-pitch relations is needed in order to supplement the present global results and to discover whether, for example, tonal differences are preserved in local detail, in addition to the global properties found in the present study. For this purpose, initially use of a correlation procedure is planned.

Again, it has to be said that since the results for Anyi were so clearly downtrend oriented, and especially since the song results were even more clearly downtrend oriented than the speech, it turned out that the use of Bete as a reference model is not strictly necessary. However, the verification of a discrete level system suggests that there may be (as in European music styles) discrete level song genres in Bete and in related languages, and that such genres would be more towards the 'natural' end of the Conventional-Natural Scale for pitch. This work is planned as a future further development, and the Bete case is included to illustrate the method.

The method is being developed further in order to specify exact tone-pitch realisation rules using the existing results as a model. The motivation, besides obtaining more detailed and precise linguistic and phonetic models, is to develop a model of sufficiently high quality to permit operationalisation in a natural and comprehensible model of speech synthesis. For this purpose, a more detailed pitch realisation relation between sequences of tone-pitch pairs is needed, in which sequential ditonal pairs H-L, L-H, H-H and L-L are modelled.

Using the kinds of analytic approach and modelling which are described in the present paper, practical speech synthesis systems for Bete and Anyi are being designed, mainly for use in the validation of annotation data using the method described by Gibbon & Bachan (2008), but also with the prospect of deployment in regional information and telecommunication systems.

REFERENCES

Agawu, V. Kofi (1988). Tone and tune: the evidence for Northern Ewe music. Africa 58.2: 127-146.

Ahoua, Firmin (1998). *Prosodic Aspects of Baule. With Special Reference to the German of Baule Speakers*. Köln: Rüdiger Köppe Verlag.

Aka, Konin (2007). *La vie musicale chez les Agni-N'dénéan (région est de la Côte d'Ivoire)*. Musée royal de l'Afrique centrale, Tervuren.

Blacking, J. (1973). *How musical is man?* Seatttle: University of Washington Press.

Connell, Bruce (2002). Dowdrift, downstep and declination. In: Gibbon, Dafydd and Ulrike Gut, eds. *Typology of African Prosodic Systems. Bielefeld: Bielefeld Occasional Papers in Typology* 1.

Connell, B. (2005). Speech and song melody in Mambila. Phonetics and Phonology in Iberia Conference.

Gibbon, Dafydd, Roger Moore & Richard Winski, eds. (1997). *Handbook of Standards and Resources for Spoken Language Systems*. Berlin: Mouton de Gruyter.

Gibbon, Dafydd, Eno–Abasi Urua & Ulrike Gut (2003). A computational model of low tones in Ibibio. *Proceedings of the International Congress of Phonetic Sciences*, Barcelona, 2003, pp. 623-626.

Gibbon, Dafydd, Jolanta Bachan (2008). An Automatic Close Copy Speech Synthesis Tool for Large-Scale Speech Corpus Evaluation. In: *Proceedings of the Sixth International Language Resources and Evaluation Conference (LREC'08)*. Paris: European Language Resources Association (ELRA).

Gippert, Jost, Nikolaus P. Himmelmann & Ulrike Mosel, eds. (2006). *Essentials of language documentation*, *Trends in Linguistics*, *Studies and Monographs* 178, Berlin: Mouton de Gruyter.

Greenberg, J. H. (1949) Hausa verse prosody. *Journal of the American Oriental Society* 69.125-135.

Hérault, Georges, ed. (1982). *Atlas des Langues Kwa de Côte d'Ivoire*. Abidjan: ILA and Agence de coopération culturelle et technique.

Hirst, Daniel & Albert di Cristo, eds. (1998). *Intonation systems. A survey of twenty languages*. Cambridge: CUP.

Kipre, Ble Francois (1989). *Morphophonologie und Morphotonologie der Nominalphrase einer Kru-Sprache (Bete der Mundart von Daloa)*. Dissertation, Universität Bayreuth.

Leben, W. (1985). On the correspondence between Linguistic Tone and Musical Melody. In Goyvaerts, Didier (ed). African Linguistics: *Essays in memory of M.W.K. Semikenke*. Amsterdam: John Benjamins, 335-343.

Leben, W. (1981). Tone in the verbal system of Anyi [sanvi]. *Proceedings of the 11th Conference on African Linguistics*.

Marchese, Lynell (1983). *Atlas linguistique kru. Essai de typologie*. Institut de Linguistique Appliquee, Universite d'Abidjan.

Richard, P. (1972). A quantitative analysis of the relationship between language tone and melody in Hausa song. *African Language Studies* XIII, 137-161.

Schneider, M. (1943-1944). Phonetische und metrische Korrelationen bei gesprochenen und gesungenen Ewe-Texten. *Archiv für vergleichende Phonetik* 7(1/2), 1-6.

Schneider, M. (1961). Tone and Tune in West African Music. *Ethnomusicology* 5:3. 204-215.

Simmons , D.C.(1980) *Extralinguistic uses of tonality in Efik folklore*. University of Alabama Press.