MODELLING SPEECH-SONG RELATIONS: AN EXPLORATORY STUDY OF PITCH CONTOURS, TONES AND PROSODIC DOMAINS IN ANYI

Dafydd Gibbon, Firmin Ahoua & Adj épol éKouam é

Universität Bielefeld, Germany; Université de Cocody, Abidjan, Côte d'Ivoire
gibbon@uni-bielefeld.de; fahoua2003@yahoo.fr

ABSTRACT

We investigate the relation between performances of speech and of song in the *awosi* dirge genre of Anyi (Kwa, Côte d'Ivoire). The relation between speech and song prosody has not been extensively studied, in terms of either structure or performance variation. We address phonetic issues of pitch height, tonal prominence, downtrends of speech in relation to melodic trends in song, and the phonological issue of prosodic domains. We propose a Conventionality Scale for speech and song, and suggest that conventional song genres in a tone language such as Anyi, with a heavily constrained prosodic system, are eminently suitable for such comparisons.

Keywords: prosody, speech, song, tone, domain

1. INTRODUCTION

Mappings between speech prosody (timing, pitch patterns) and the rhythm and melody of songs vary from culture to culture and style to style within cultures. We suggest that mapping may be located along a linear scale from the 'conventional' (musical autonomy) to the 'natural' (music-speech correspondence), and investigate prosodic parameters in the *awosi* (dirge) song genre in Anyi (Kwa, Tano, Bia; ISO 639-3: any; Côte d'Ivoire) in the Sanvi dialect of Aboisso. The investigators comprise two linguists who are familiar with the Anyi language, and a musicologist who is a native speaker and expert performer of Anyi music.

Morpho-lexical functions of tones, prosodic constraints on tones and tonal domains have been the focus of many studies. Pitch downtrends are increasingly well understood [5]. But it is still unclear how these features affect songs, despite the informative studies by Connell [4] Greenberg [6], Richards [8], Schneider [9, 10], Simmons [11], which are only partly comparable because they deal with different genres. Leben [7] has argued that speech tones and song tunes in Hausa are related: lexical tones and downdrift are well preserved, and thus the comprehension of songs should not be affected. Blacking [3] observes a close correspondence between tones in speech and in songs in Venda, similarly Connell [4] in Mambila. Agawu [1] finds that music is less constrained by speech in Ewe.

We select the *awosi* genre partly because of an initial intuition that this genre shows some speech-like characteristics, unlike more formal ceremonial genres, but also partly because for obvious natural reasons dirges are performed quite frequently, unlike songs for seasonal or courtly festivities.

In our preliminary study of speech-song relations, we use a qualitative methodology with traditional auditory phonetic analysis, illustrated with F0 traces and a transcription in conventional European stave notation, and discuss two prosodic parameter types: pitch contours and prosodic domains. A data-intensive quantitative analysis is not performed, because (1) it will never be possible to find sufficient data for a quantitatively interesting analysis because of the restricted social function of the genre, the restricted inventory of such songs, and strict cultural constraints on repeating the songs out of context; (2) a basic qualitative methodology is needed for cooperative work between linguists and musicologists; (3) currently the political situation in the region unfortunately does not permit further fieldwork.

2. HYPOTHESIS; DATA; SOUND SYSTEM

2.1. Hypothesis

In order to accommodate the complexities of speech-song interaction, for present purposes we simplify the multidimensional matrix of possible similarities and differences to a linear Conventionality Scale (Figure 1), from 'stylised' melody and timing in music to complex natural determinants of melody and timing in speech. We hypothesise that more natural musical genres have fewer stylised and more speech-like features, and

that styles with both ceremonial and emotional content such as *awosi* will occupy an intermediate position on the scale.

Figure 1: Music-speech Conventionality Scale for prosody.

conventional	*natural*
music-like	*speech-like*

2.2. Data

The song recordings were made around 1970 for a television production; very few are available. Following Connell [4], male and female spoken versions of the texts of the songs were made; the original singer was unfortunately no longer available. We illustrate our analysis here with a couplet in which there is a repetition of the following two-sentence sequence:

mí	ní	ʤʉ	à	Wú	ó
my	*mother*	*-topic*	*past*	*die*	*emph*

'my own mother has died'

è	ɟɔ́	èɟɔ́	bɔ́	ngɑ́	ngèlé	wɑ́
matter	that	past	to me	I shall	tell	whom

'As to what I have been thinking of, whom should I tell?'

2.3. Sound system

The minimal detail supplied here is sufficient for the present purpose. Burmeister [2] provides more detail. The Anyi vowel and consonant phoneme spaces are shown in Figure 2 and Figure 3.

Figure 2: Anyi vowel space; each vowel may be -ATR or +ATR, yielding a total of 16 vowels.

i ĩ		u ũ
e		o
a ã		

Figure 3: Anyi consonant space.

p	t	c	k	kp	
b	d	ɟ	g	gb	
f	s	ɕ			H
m	n				
	l	j	w		

Anyi-Sanvi syllable structures are V /a, ɔ, ɛ/ (affix), N (homorganic nasal prefix), CV (all vowels), with ATR vowel harmony. There are two lexicalised contrastive tones, H and L (marked with acute and grave diacritics), with complex realisations due to spreading as HL, LH (cf. Figure 4) downstep; cf. Connell [5]. Same-tone sequences remain nearly level and automatic downstep occurs, (the H-L fall in pitch is greater than the L-H rise, yielding a 'terraced' pitch downtrend).

Figure 4: Anyi tone system (surface tones).

fɔ̀ʊ̀	Touchy	fɔ́ʊ́	right		
ʤʉ̀é	Food	ʤʉ́è	day		
bɔ̀lè	potash, venom	bɔ̀lé	fog	bólè	packet
èsɛ́	Then	èsɛ̂	pot		
Èbô	River	èbǒ	field		

3. SPEECH-SONG COMPARISON

3.1. Comparison of pitch contours

We note several syllable modifications (e.g. contractions, elisions), but do not treat them here.

We proceed as follows:

(1) Establishment of underlying and surface tones of the utterance. These are shown in the following example with the underlying tones above the surface tones (note the rule which spreads H on to a following L yielding surface HL):

mí	ní	ʤʉ	à	wú	ó
H	H	L	L	H	H
H	H	HL	L	H	H

è	ɟɔ́	èɟɔ́	bɔ́	ngɑ́	ngèlé	wɑ́
L	H	LH	H	H	LH	H
HL	H	H.LH	H	H	HLH	H

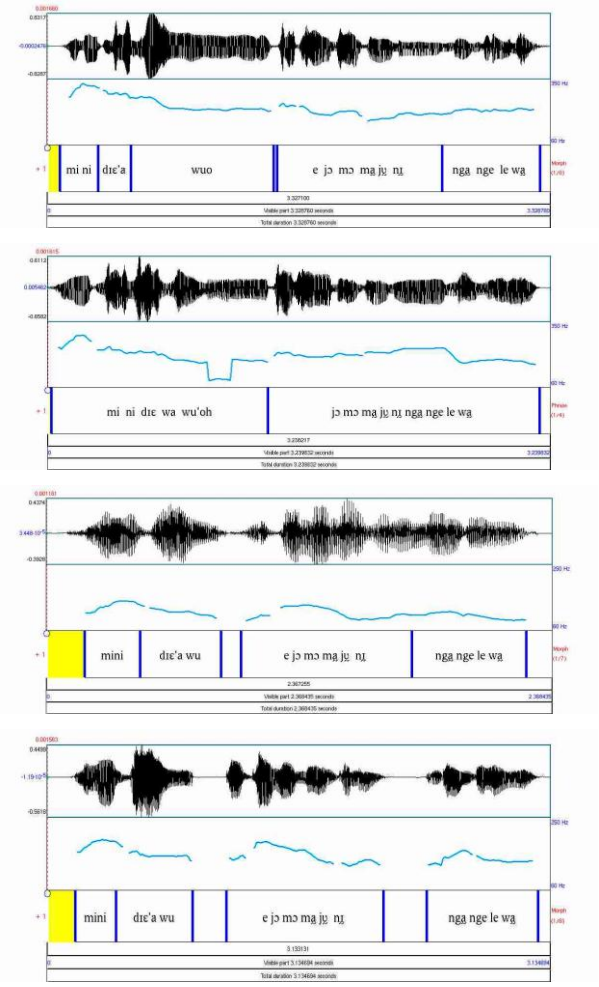
(2) Description and comparison of the pitch contours of spoken and sung versions by conventional perceptual analysis, clarification with F0 traces.

The four pitch traces in Figure 5 show the first and second lines of the song (the second being a repetition of the first), and male and female read speech versions. The sharp dip in the second trace is due to creaky voice associated with low pitch; the gaps in the third and fourth traces are pauses. All traces are downtrending. The pitch in the first trace (first line of the couplet) ends in a speech-like sustained pitch, but all others have final lowering. The H tones are observed in all contours.

The sung version is more 'stylised', having an overall smoother contour, i.e. less abrupt pitch variation. The sung rendering has much higher minimum, maximum and mean pitch (min-max-mean Hz: 154-310-215) in relation to male read speech voices (e.g. 87-158-117, 77-146-106); in

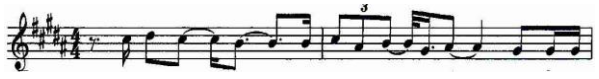
the song the pitch range is about 1 octave, in the spoken version slightly less.

Figure 5: *Awosi* song (top 2 traces), male and female readings.



The song was also transcribed in conventional European stave notation, with simplifications constrained by the notation (cf. Figure 6 for the second line). The stave notation reflects the main relational properties of the pitch contour as shown by the pitch trace; however the notes in the stave notation are not to be taken to reflect exact absolute pitches and frequency intervals or exact durations.

Figure 6: Stave notation of awosi song (repetition).



3.2. Comparison of prosodic domains

The surface prosodic domains for spoken and sung versions were established empirically from temporal phrasing criteria (rhythm, length, pause):

Song:
(mí ní đɪ à wú ó)(è ɟɔ́ bɔ́ mą́ ɟų̀ nį́) (ngą́ ngè lé wą́)

Speech:
(mí ní đɪ à wú ó)(è ɟɔ́ bɔ́)(mą́ ɟų̀ nį́) (ngą́ ngè lé wą́)

Of the three domains in the song, and the four domains in the speech, the first is identical in each case, and so is the last. However, the single central domain in the song corresponds to two domains in the speech: the subject phrase /è ɟɔ́ bɔ́/ and the predicate phrase /mą́ ɟų̀ nį́/ have independent domains, while both constituents cohere in a single prosodic domain in the song.

Whether this is due to individual variation on the part of the speaker in the speech rendering, or a symptom of a consistent tendency in speech as opposed to song is not clear from the restricted data set. Two generalisations are suggested: (1) the 'same prosody' null hypothesis at the 'natural' end of the Conventionality Scale is (tentatively) falsified in view of the stylisation and domain differences; (2) the speech-like prosodic properties are simultaneously confirmed in view of the coincidence of prominences and downtrending. These two properties indicate that *awosi* takes up an intermediate position on the Conventionality Scale, with stylised properties, but at the same time also with natural speech properties.

4. CONCLUSION

We have outlined a procedure for comparing a restricted set of parameters of speech and song, focussing on the *awosi* genre of the tone language Anyi, and looking at two prosodic parameter types: pitch contour and prosodic domain. Both of these are seen in a qualitative auditory phonetic rather than a formal phonological or a data-intensive quantitative perspective.

The Conventionality Scale which we introduce turns out to be useful in accounting even for small speech-song differences of the kind which we have observed in the data: at the speech end of the scale, pitch changes are more abrupt and contours less holistic, than toward the music end of the scale.

The *awosi* dirge genre can be characterised as intermediate on the scale because of the presence of prosodic similarities such as downtrend and H tone prominence on the one hand, and differences in stylisation (contour smoothing) and grammar constraints on domain size for speech on the other. We also note that the speech-song similarities tend to corroborate the results of Leben [7] and Connell [4], at least for the Anyi *awosi* genre. With our

cautionary remarks on data set size in mind, we present these results as sources of hypotheses for further work, and the present qualitative auditory study has indeed provided a new and useful foundation for ongoing and future more detailed studies, including quantitative acoustic phonetic analyses along the lines of Connell [4] of more of the available material.

On the theoretical plane, our study suggests that there is a set of well-defined prosodic parameters which need to be considered when positioning speech and song in terms of their differences and similarities. The parameters we have looked at are:

1. Pitch:
    1. Overall pitch height and range.
    2. Downtrend (including downstep).
    3. Local tonal prominences.
2. Prosodic domains:
    1. Long grammatical domains (sentence).
    2. Small grammatical domains (subject phrase, predicate phrase, phonological word).

There are many relevant parameters which we did not discuss in the present context, including prosody-segment relations such as syllabification, syllable lengthening and reduction, and fast-speech or musically determined assimilation and rhetorical (final 'o') and metrical vowel lengthening. The relatively large parameter set suggests that a more highly dimensional 'Conventionality Space' would be appropriate in future research than a linear Conventionality Scale.

In view of the overlap of prosodic properties between speech and song, we speculate that speech-song relations are a special case of style variation (cf. chant-like listing pitch contours in English), and therefore that speech and song are structurally comparable, and not fundamentally different. The small data set can clearly only provide a first approximation, but we suggest that our results are compatible with at least some of the results of earlier studies, and that the concepts of Conventionality Scale and Conventionality Space open up new avenues of investigation for studies in other languages.

5. ACKNOWLEDGMENTS

We are deeply indebted to J. Burmeister, B. Connell, L. Hyman, W. Leben, F. Oyebade and E.-A. Urua for their insightful comments on an earlier version.

6. REFERENCES

[1] Agawu, V.K. 1988. Tone and tune: the evidence for Northern Ewe music. *Africa* 58(2), 127-146.

[2] Blacking, J. 1973. *How Musical is Man?* Seatttle: University of Washington Press.

[3] Burmeister, J. 1982. Agni. In Hérault, G. (ed.), *Atlas des Languages Kwa*. Abidjan: Institut de Linguistique Appliquée.

[4] Connell, B. 2005. *Speech and Song Melody in Mambila.* Lecture Notes, Barcelona.

[5] Connell, B. 2011. Downstep. In van Oostendorp, M., et al. (eds.), *The Blackwell Companion to Phonology 5*. Malden MA & Oxford: Wiley-Blackwell, 824-847.

[6] Greenberg, J.H. 1949. Hausa verse prosody. *Journal of the American Oriental Society* 69,125-135.

[7] Leben, W. 1985. On the correspondence between linguistic tone and musical melody. In Goyvaerts, D. (ed.), *African Linguistics: Essays in Memory of M.W.K. Semikenke*. Amsterdam: John Benjamins, 335-343.

[8] Richards, P. 1972. A quantitative analysis of the relationship between language tone and melody in Hausa song. *African Language Studies XIII*. 137-161.

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

[10] Schneider, M. 1961. Tone and tune in West African music. *Ethnomusicology* 5(3), 204-215.

[11] Simmons, D.C. 1980. *Extralinguistic Uses of Tonality in Efik Folklore*. University of Alabama Press.