Tone and intonation

Bernard Caron
Shlomo Izre'el

Most of the literature on intonation derives from pioneering studies on English intonation. These authors and their followers have identified the exponents of intonation as pitch, rhythm (including length and pauses) and intensity. The difficulty when studying intonation in "tone languages" is that pitch is already mobilised by the lexicon and the morpho-syntax[1]. The question is then: does pitch play a role in the intonation of tone languages, and how? Is this role comparable to that of pitch in non-tonal languages? This problem became crucial in the transcription and segmentation of the tonal languages represented in the CorpAfroas corpus of Afroasiatic languages, *viz*. Hausa and Zaar, two Chadic languages on the one hand, and Wolaitta, an Omotic language on the other hand.

The objective of this study is to identify the basic components of pitch that can be isolated from tone and attributed to intonation, and establish them as the elements that must be accounted for in the transcription of an oral corpus in order to make it useful for typological studies of intonation.

To address this problem, this study will lean heavily on Zaar, a Chadic tone language spoken in the South of Bauchi State, Nigeria.

I. Introduction

Zaar is a tone language with five phonemic tones, 3 level tones : High (written with an acute accent: **á**), Mid (left unwritten: **a**) and Low (written with a grave accent: **à**); and 2 contour tones: Falling (written with a circumflex accent: **â**) and Rising (written with a caron: **ǎ**).

Tones are important to identify lexemes (lexical tone), but play a part in morphosyntax too. The influence of morphosyntax on the surface realisation of lexical tones is explained e.g. in the suprasegmental theory of tone by post-lexical rules[2].

Let us take the verbs **vər,** give and **ʧet,** tell. Tense, aspect and mood are expressed in Zaar by "subject pronouns" which can affect the tone of the following verb. If we take the subject pronouns **tà**:, 3PL.PFV and **tə́,** 3PL.AOR, and combine them with these verbs, we get the following sentences, where the tones of the verbs vary according to their lexical tone classes and the subject pronouns preceding them:

[1] Tone-languages use pitch variation for morpho-syntactic uses on a scale that cannot be compared with what obtains in a stress-language like English for example, where stress combines variation in pitch, loudness and length in the same way as intonation, but where the often quoted opposition between verb and noun relying on stress (e.g. *a/to record* = [ˈrekəd]/[riˈkoːd]) is at best marginal and not systematic (e.g. *a/to cover* = [ˈkɔvə] in both cases). The only place where stress manifests itself as a prominent grammatical feature, concerns its association with certain suffixes, e.g. *–ics*, *-aphy*, *-ition*, etc., and it is not distinctive.

[2] The tone marking in the transcription has been done by Marvellous S. Davan, a language assistant who consistently marks postlexical tones in grammatical words, whatever their acoustic realisation. Some of the recordings of the interviews used for this study were made in 1999. At that time, Zaar was a language without grammatical tradition or orthography, and up till now, it is not taught at school. The transcriber had just been trained in marking lexical tones on individual words and was given the recordings to transcribe as a first exercise. Without hesitation, he marked post-lexical tones right from the beginning, and has never varied in his transcriptions. Some passages that had not been faithfully transcribed (the hesitations, syntactic mistakes, etc. had been eliminated in the original transcription) were done again 10 years later. The same passages were transcribed again with exactly the same tones.

- **tà**: **vər**, 'they have given (it)' ; **tə́ vàr**, 'they gave (it)'
- **tà**: **ʧet**, 'they have said (it)' ; **tə́ ʧêt**, 'they said (it)'

Verbs can affect the tone of direct object pronouns through tone spreading, as can be seen in the 2nd person singular direct object pronoun **=kə** in the following sentences:

- **tà**: **vər=ɣə**[3] , 'they have given you' ; **tə́ vàr=ɣə̀**, 'they gave you';
- **tà**: **ʧtet=kə**, 'they have told you' ; **tə́ ʧét=kə̀**, 'they told you.'

However, these "surface tones" accounted for and/or predicted by post-lexical tonological rules undergo further variations. This can be heard when listening to recordings of natural speech, and it can be represented and measured through instrumental acoustics. Our hypothesis is that the pitch component of intonation in tone languages lies in this variation.

A. Tone variation

Pitch varies all along the syllables. Even non-modular tones are rarely realized by a plateau. The measurement of pitch in the study of tone has recently used the notion of "target", defined as follows by (Akinlabi & Liberman 2001) : "the phonetic target value of a tone [is] the highest F0 of a High tone, or the lowest F0 of a Low tone ". In Yoruba, this target "is found at the end of the span of time corresponding to the associated tone-bearing unit". For Zaar, this is true for Low and Mid tones, but not for for High tones, where the phonetic target is at the beginning of the tone bearing unit. This method is fastidious, repetitive, time-consuming, but most of all, very much dependant on the researcher's judgement and intuition as to where the actual target is located. And this judgement may be biased and subject to questioning.

Another way, which we have used for this study, is to transcribe prosody using pitch contour stylization based on a tonal perception model and automatic segmentation, as done e.g. by "*Prosogram*". The system has been implemented by Piet Mertens as a Praat[4] script. *Prosogram* follows four steps:

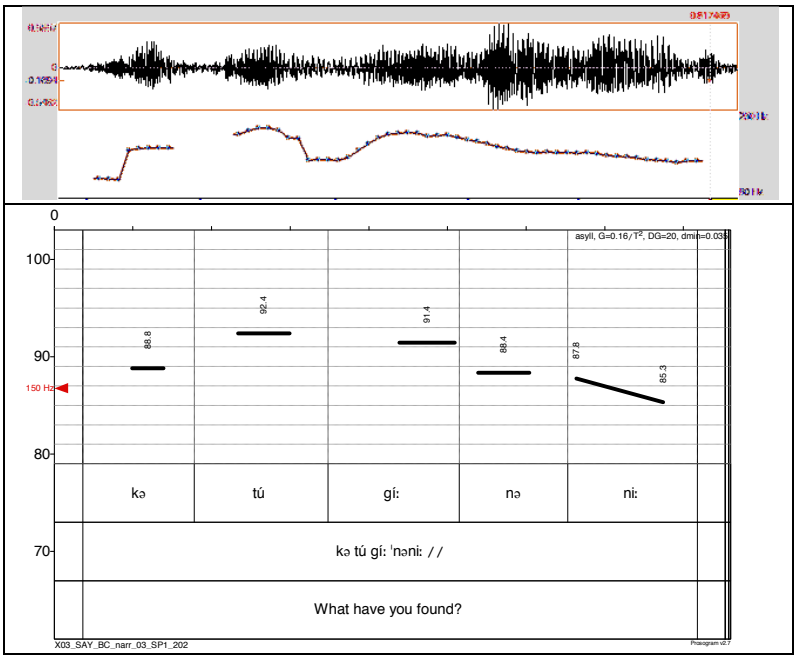
- Calculate acoustic parameters: F0, intensity, voicing (V/UV).
- Obtain a segmentation into units of the types indicated above. Select the relevant units (e.g. vowels, syllables). Select the voiced portion of these units, that has sufficient intensity/loudness (using difference thresholds relative to the local peak).
- Stylize the F0 of the selected time intervals.
- Determine pitch range used in speech fragment. Plot stylized pitch and some annotation tiers (text, phonetic transcription). Use a musical (semitone) scale and add calibration lines at every 2 ST for easy interpretation of pitch intervals[5].

[3] /k/ is realised as [ɣ] in the contexts V_V and r_V.
[4] "Praat"is a tool for acoustic and phonetic research, written by Paul Boersma and David Weenink, of the Institute of Phonetic Sciences in Amsterdam.
[5] The analysis of pitch intervals done by Prosogram is based on the glissando threshold G, or auditory threshold for pitch variation. This depends on the amplitude (extent) and the duration of the F0 variation. Since the work of J. 't Hart, it is usually expressed in ST/s (semitones per second). ST use a logarithmic scale to give a better approximation of the way F0 is perceived and interpreted by the human ear. For convenience reasons, the 'Automatic detection of syllabic nuclei' has been selected for this work in the Prosogram settings. This method uses "a segmentation into local peaks in the intensity of band-pass (300-3500 Hz) filtered speech, adjusted on the basis of the intensity (full bandwith)". The other methods require a manual time-plotting of the syllable

The advantage of this method is twofold: first, it provides in a simple way images of Intonation Units, plotted against a time scale and customized annotation tiers; second, as it is automatic, it is reproductible and objective.

In the presentation of our examples, the representation of prosody as calculated by Prosogram is plotted against three tiers (1) the transcription of the syllables isolated by Prosogram ; (2) the transcription of the Zaar text with post-lexical tones and prosodic annotations; (3) a free translation into English. In the ex. below, the Prosogram stylization is preceded by the pitch curve as represented by Praat.

Example 1 (X03_SAY_BC_narr_03_SP1_202_001)



Our hypothesis is that the role of pitch in Zaar intonation can be observed in the variation between post-lexical tones as they are perceived and transcribed by the native speaker and their acoustic realisation as represented by Prosogram. These variations, i.e. the way intonation influences the realisation of post-lexical tones, fall under the following categories :

(a) Declination;
(b) Intonemes[6], which are divided into:
- Final intonemes : Fall (↓), Rise (↑), Level (→) and Rise-Fall (↑↓);
- Initial intonemes: Downstep (!) and Upstep (¡).

In the first part of the paper, we will illustrate these prosodic features (declination and intonemes). Then the final part, we will analyse a few intonation patterns exemplifying the combination of these features. But before doing this, let us clarify some terminological issues.

---

nuclei on a Praat Textgrid. [http://bach.arts.kuleuven.be/pmertens/prosogram/#Rationale], consulted 07/02/2011.

[6] Intonation literature uses the terms "falling/rising tone" to refer to phonetic cues characterised by an increase/reduction in pitch. In the linguistic description of tone languages, "tone" refers to the pitch variations used to characterize lexical and grammatical oppositions alongside vowels and consonants. We want to preserve the terms "falling/rising tone" for these phonologically distinctive suprasegmental units. For the increase/reduction in pitch working as intonation acoustic cues, we will use the terms "Fall/Rise intoneme" rather than tone.

B. Intonation units

Intonation units

The minimal Intonation Unit is what "encapsulates a functional, coherent segmental unit, be it syntactic, semantic, informational, or the like" (Izre'el & Mettouchi, in this vol.). In other words, it is "that part of a discourse text that the speaker by his voice wished to identify as an informational unit." (Markus 2006:112)

An intonation unit (henceforth IU) is characterized by a combination of the following elements :

- overall : declination.
- final : pause, creaky voice; lengthening of final vowel or consonant;
- initial : (upward or downward) pitch reset, acceleration.

Paratones

The paratone corresponds to a functionally complete speech act[7]. By default, it is followed by a pause, ends in a Fall, and is characterized by overall declination.

Major and minor intonation units

Paratones can be simple or complex. A complex paratone consists of one or more minor units and finishes with a major IU. A simple paratone consists of a single major IU. Apart from the final Level intoneme specific to the minor unit, phonetic cues associated with the end of minor units are not much different from those of major IU. They are: pause, Rise, or Fall, lengthening of final vowel or consonant, etc., and are followed by pitch reset. What distinguishes minor from major IU is the fact that they do not correspond to the end of a complete speech act. The end of a minor IU is transcribed by a single slash (/).The end of a major IU is transcribed by a double slash (//) corresponding to the completion of the speech act. In a transcribed text, a paratone is delimited by two doubles slashes: it ends in a double slash and begins after the final double slash of the preceding paratone.

A period is the highest prosodic hierarchy, defined as "a speech stretch that shows declination along its paratones ('supradeclination' according to Wichman 2000: §5.2.2)" (Izre'el & Mettouchi, in this vol.)

II. Declination

For both tone and non-tone languages, declination has been presented as a universal tendency due to physiological constraints[8], linked to the energy used to expel pulmonic air through the vocal organs. This creates the background for a "neutral" intonation against which variations of pitch by the speaker can be interpreted as meaningful patterns of deviations[9].

---

[7] Speech act : declaration (positive and negative), question (including rhetorical questions), injunction, exclamation, etc.

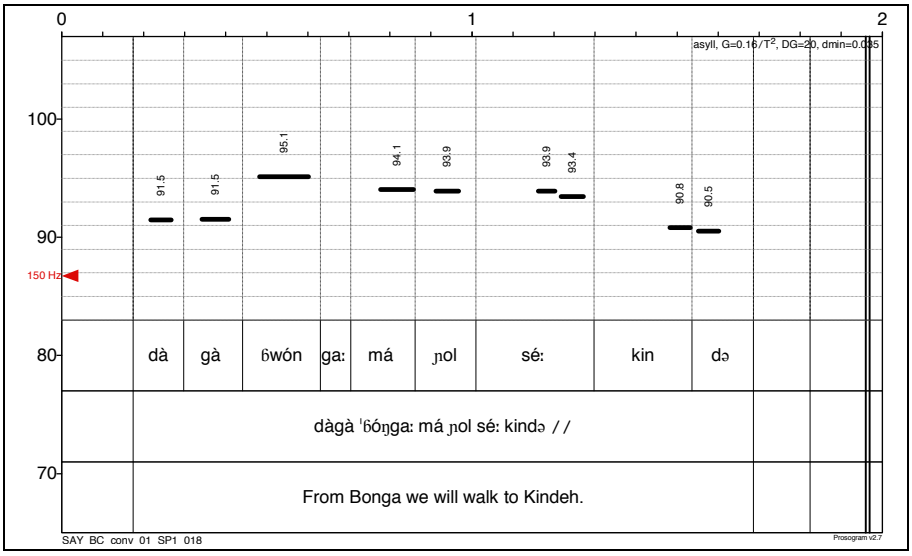
[8] " (…) F_0 tends to decline over the course of phrases and utterances, both in tone languages and in languages like English or Dutch." (Ladd 1996:73ff., quoted in Bearth 1998: n. 2)

[9] However, Bearth (1998) presents data from Toura, a four-tone African language, where declination is limited to local tonal downstep where two tones of the same phonological level are separated by one or several lower tones, the second tone tends to be realised lower than the first. This lowering is then immediately locally readjusted and the following tones resume the general framework of the language, where all high tones of a unit will be pronounced at approximately the same level. Intonation is then expressed at the periphery of the IUs. "C'est la dernière more (...) de l'énoncé qui est le point de contact entre la [tonalité lexico-grammaticale et l'intonation périphérique] à partir duquel se contitue un paradigme énonciatif chargé de caractériser l'énoncé des points de vue notamment de son statut en tant qu'acte illocutif, de sa complétude ou incomplétude, de

How is this compatible with what obtains in tone languages where the constains of lexico-grammatical tones may influence the melody in an upward movement, contrary to the general downward movement of declination? Bearth (98) sketches a typology of declination in tone languages, falling into 3 categories illustrated by Chinese (where intonation is superimposed on lexico-grammartical tonology), Akan languages (where declination is phonologised into tone downstep, and intonation is added to the periphery) and Toura (where declination is neutralised, and intonation is added to the periphery).

In Zaar, where declination has not been phonologised into tone downstep, it can still be be observed within each intonation unit, from the minor unit up to the period, and can be observed as :a gradual lowering of the pitch over the intonation unit. This is noticeable sp. in High tones. The relative height of tones within an IU is related to stress. Intensity stress in Zaar is used to underline the rhematic status of lexemes. Stress overrides declination and gives a tone its full lexical value. The highest tone in an IU is the first High tone of the first rhematic lexeme. If this rhematic lexeme does not bear a high tone, or if the IU has no rhematic lexeme, the highest tone is the first High tone of the unit. Within that framework, each High tone following the nuclear stress of an intonation unit is pronounced lower than the following one. In example (2) below, the 1st High tone corresponds to the nuclear stress. The 3 High tones of the IU measure at 95,1, 94,1 and 93,9 respectively. Moreover, the last Mid tone of the IU is lower by 1 ST than the two initial Low tones.

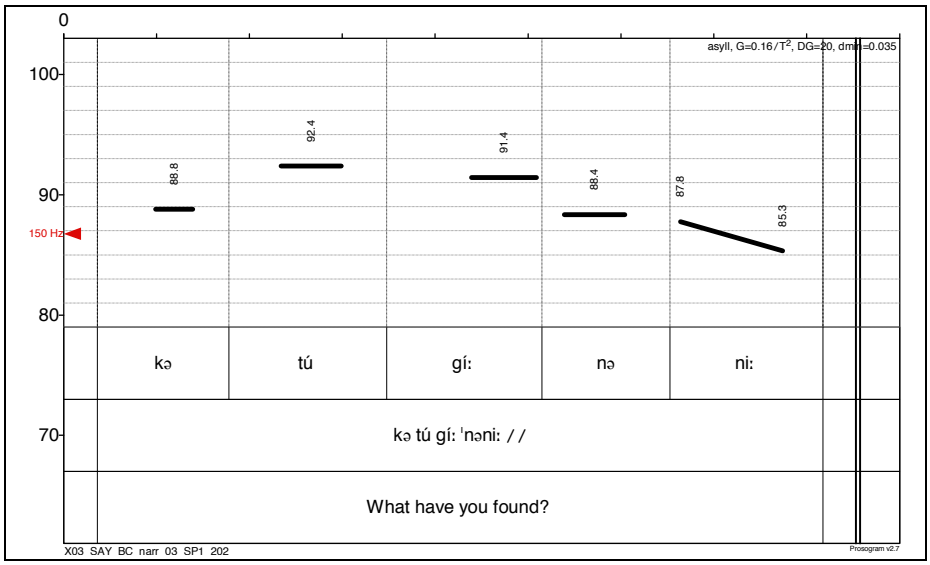
Example 2. Say_bc_conv_01_SP1_018



This makes for the canonical declarative intonation of Zaar. The same intonation pattern is found in WH-Questions, as in the following example.

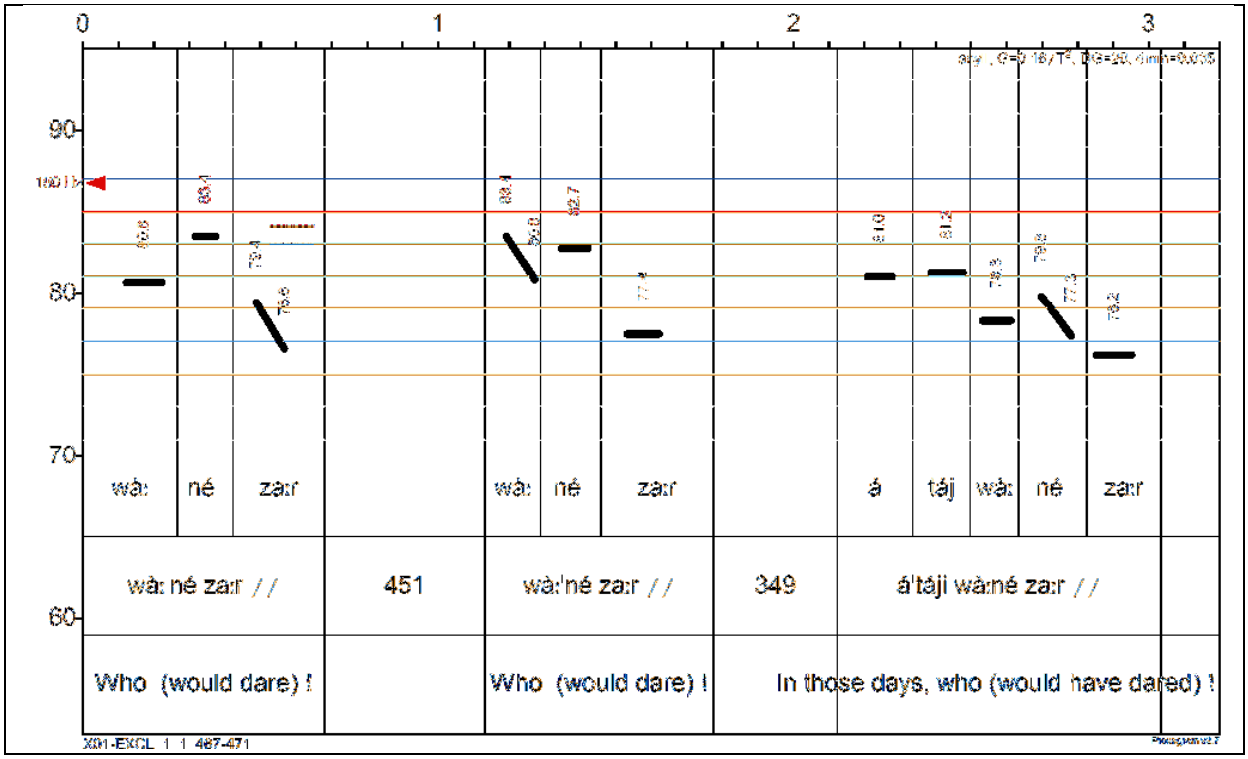
l'expression de la subjectivité et de l'émotivité, ainsi que du positionnement social des interlocuteurs. » (Bearth 1998: 80-1)

Example 3Say_bc_narr_03_SP1_202



To compensate for declination, each IU starts with an initial pitch reset. When IU's are integrated into a paratone, or paratones in a period, declination applies inside this larger unit (paratone or period): each intonation unit starts with a lower pitch reset than the preceding one and sees its High tones decline in the same way as within a minor or major unit. See below the example of a period with High tones measured at 83,4 – 82,7 – 81,2 – 79,8.

Example 4. X01-EXCL_1_1_467-471_001



Against this background of declination, which helps identify the limit of speech units through pitch reset, intonemes operate both at the initial of IUs (affecting the whole of the unit) and at the end, in what Bearth (1998) calls 'peripheral intonation'.

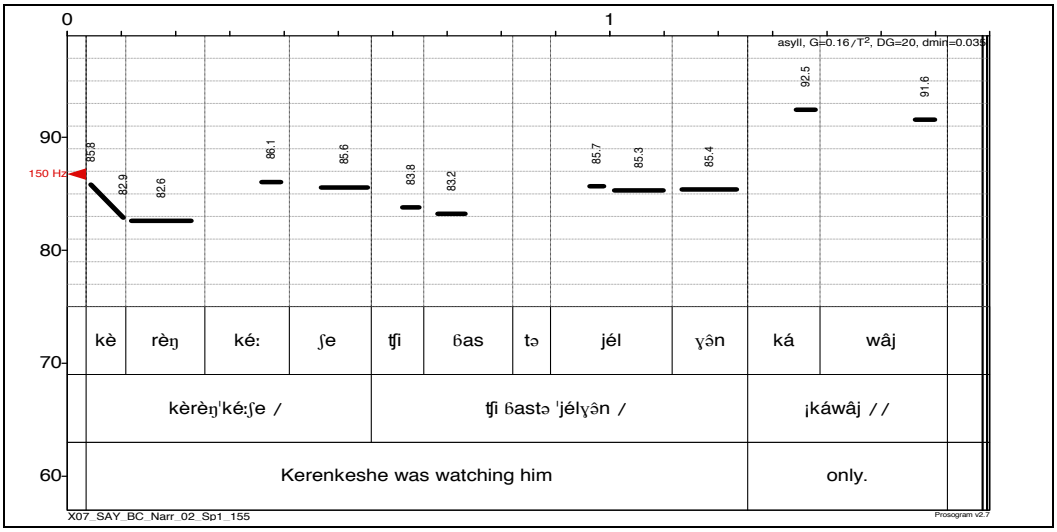
III. Intonemes

- **Initial : downstep and upstep**

These consist in a noticeable change in the register of an intonation unit compared to the preceding one. This initial lowering (Downstep, noted !) or raising (Upstep, noted ¡), creates a break in the gradual lowering of the pitch induced by declination. Both upstep and downstep are associated with specific functions: upstep is associated with topicalisation, Y/N questions, emphasis of adverbials and emotional statements. Downstep is associated with parenthesis and comments following an (upstepped) topic.

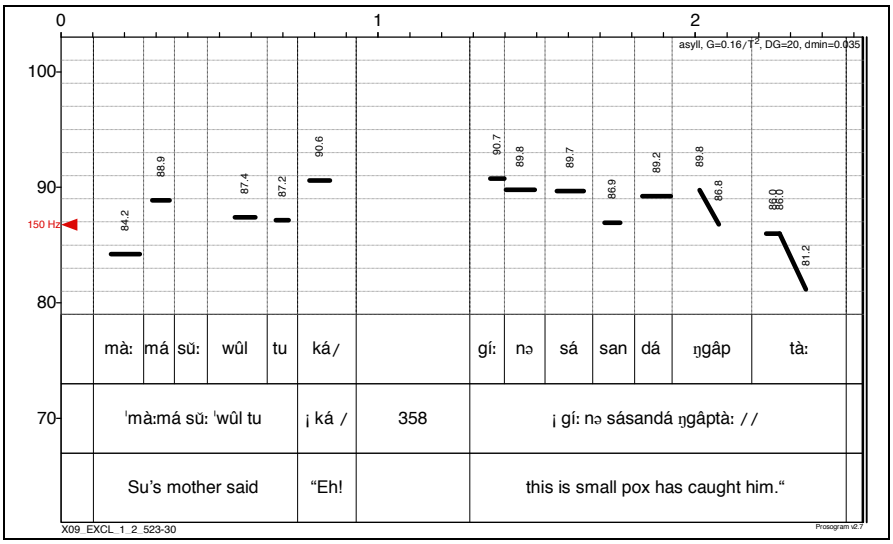
The following example shows an upstep of 7 ST emphasizing the final adverb *kawai*,

Example 5. X07_SAY_BC_Nrr_02_Sp1_155;



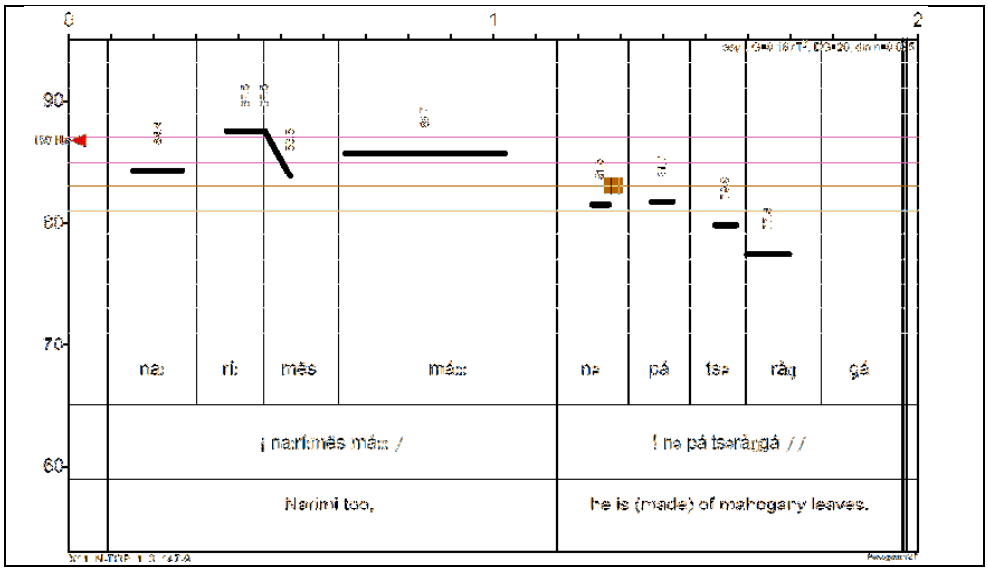
The following example shows an upstep of 3,5 ST on a reported speech concerning a very emotional passage of narration where the speaker narrates the onset of an epidemic that claimed many casualties in the children of the village:

Example 6. Emotion : EXCL_1_2_523-30



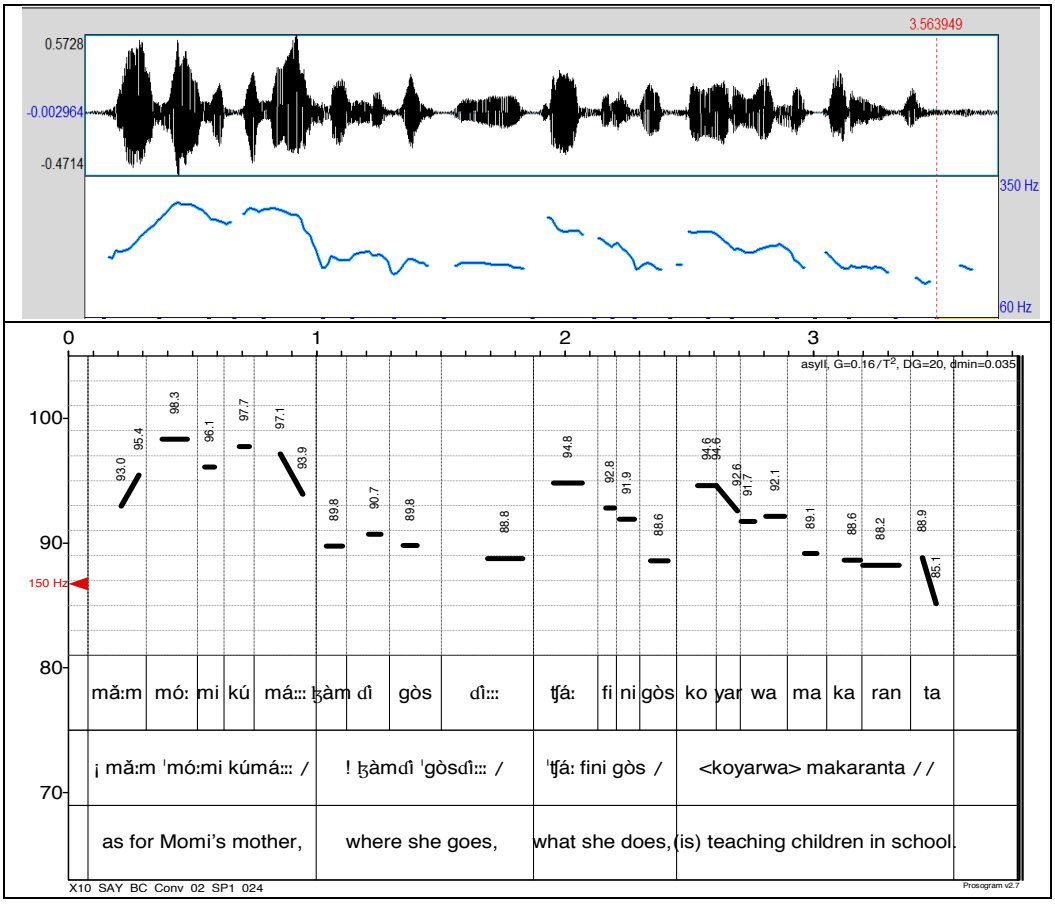
In the following example, an initial upstep contrasting a new with a preceding Topic, is followed by a downstep introducing the Comment (¡___→/!___↓//):

Example 7. N-TOP_1_3_147



In the following example, after an initial upstep corresponding to the introduction of a new Topic (a new example to prove the speaker's case), a downstep accompanies some backgrounded elements where the speaker reminds her audience of the theme of the conversation (women keep running about, overworking themselves, whereas men stay idle in the compound, chatting with their friends). This long paratone is characterized by ample declination and clear initial pitch reset at the beginning of IUs.

Example 8. SAY_BC_Conv_02_SP1_024



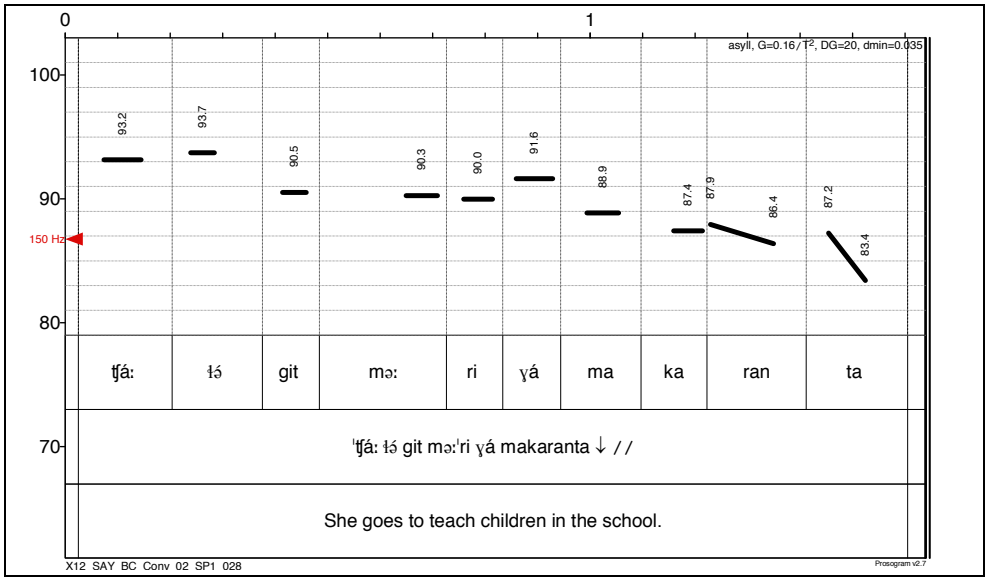
C. Final intonemes

These final (peripheral) intonemes are the Fall, the Rise, the Level, and the High-Fall.

i. Fall

This Intoneme consists in a distinctive lowering of the pitch at the end of the IU. In Prosogram, it is represented by a downward slant. This intonation corresponds to the canonical assertion and WH-questions. We transcribe it with the sign "↓" in our annotation.

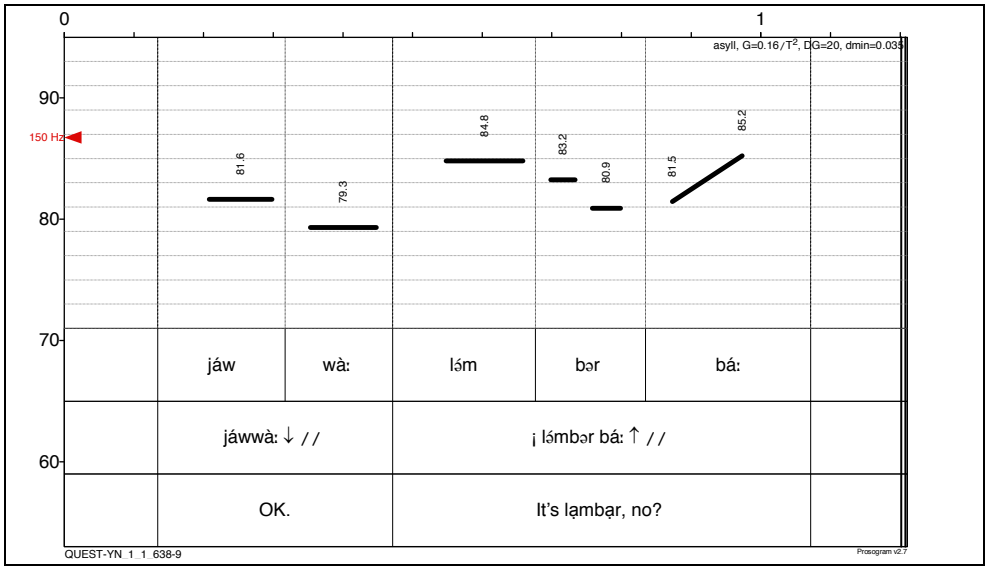
Example 9. SAY_BC_Conv_02_SP1_028



ii. Rise

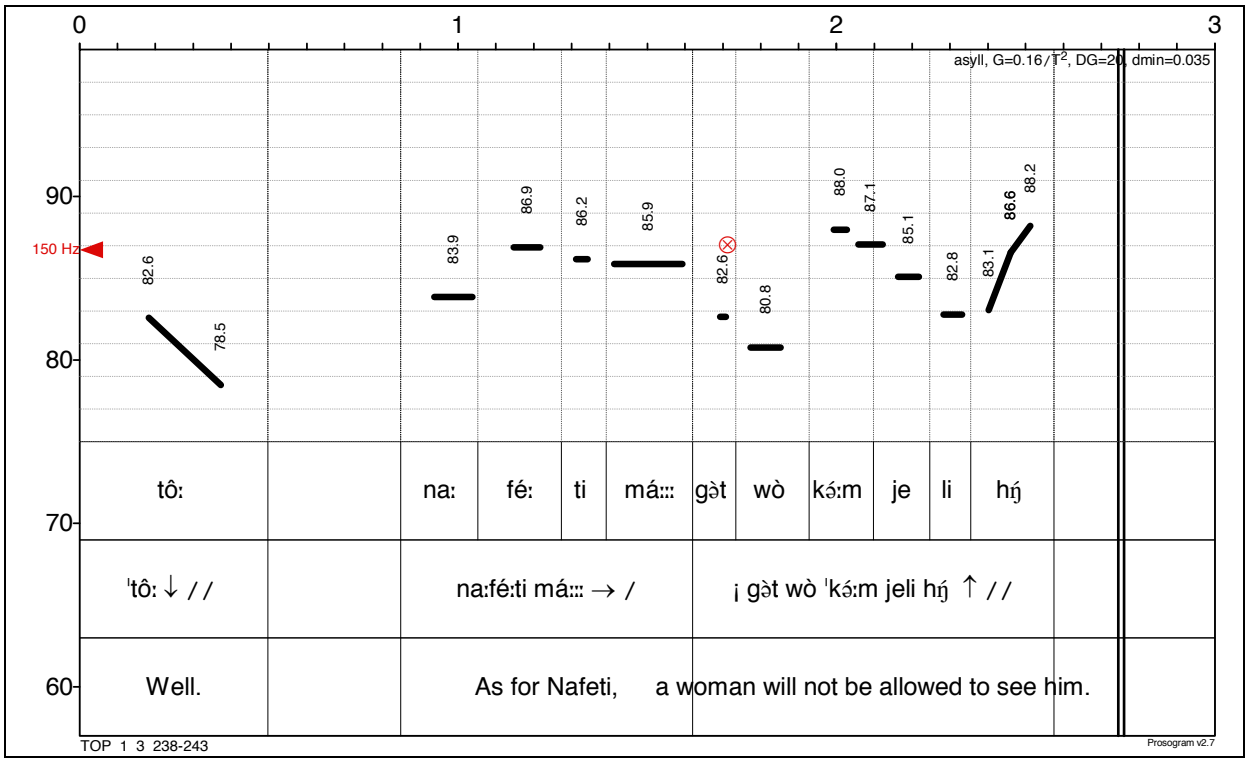
This final intoneme (transcribed ↑) is mostly associated with Y/N questions:

Example 10. QUEST-YN_1_1_638-9



It can also be associated with exclamation, such as can be seen on the final negation of the following example[10]:

Example 11. TOP_1_3_238_243



iii. Level

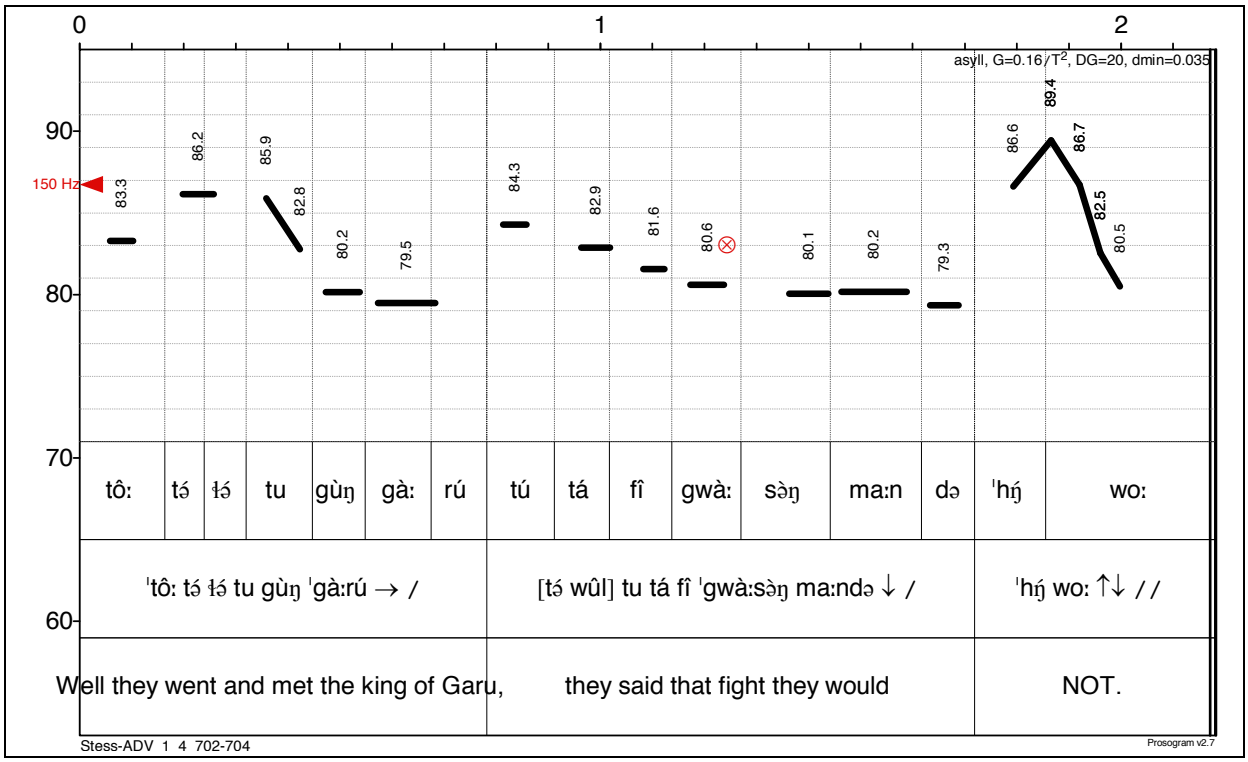
This final intoneme (transcribed →) cancels, declination. It is often associated with lengthening and induces the only (rare) cases of plateau realization of flat tones. Seen only at the end of minor units, it is often associated with topicalizing morphemes. In Prosogram, it appears as a horizontal bar. Such an intoneme can be seen in the previous example at the end of the second IU.

iv. Rise-Fall

It appears in the prosogrammes as a long bar slanting sharply downward, preceded by a shorter upward slanting or vertical bar. It is systematically associated with emphasis on negation; ideophones and assertion particles. In the following example, it falls on the final negation followed by the assertive particle ***–o:***.

[10] The low register of the first IU is due to the fact that this IU is the final major unit of a period. The upstep in the 3rd and last IU is observed on the 1st High tone of the verb ***ká:m***, "can", measured at 88 ST, compared to 86,9 ST for the highest tone of the 2nd IU. This upstep is part of the exclamative intonation.

Example 12. Stess-ADV_1_4_702-704_001

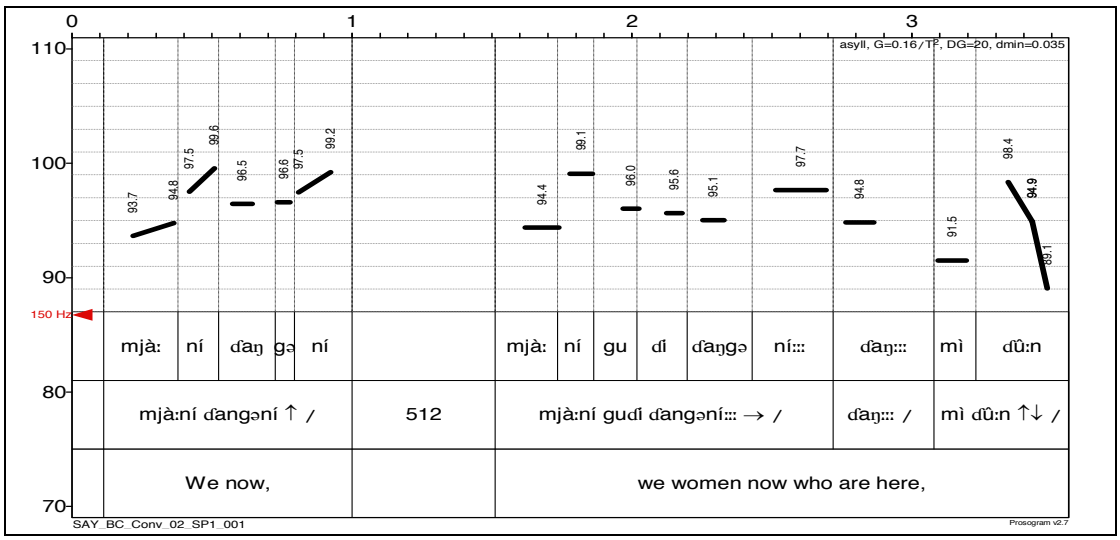


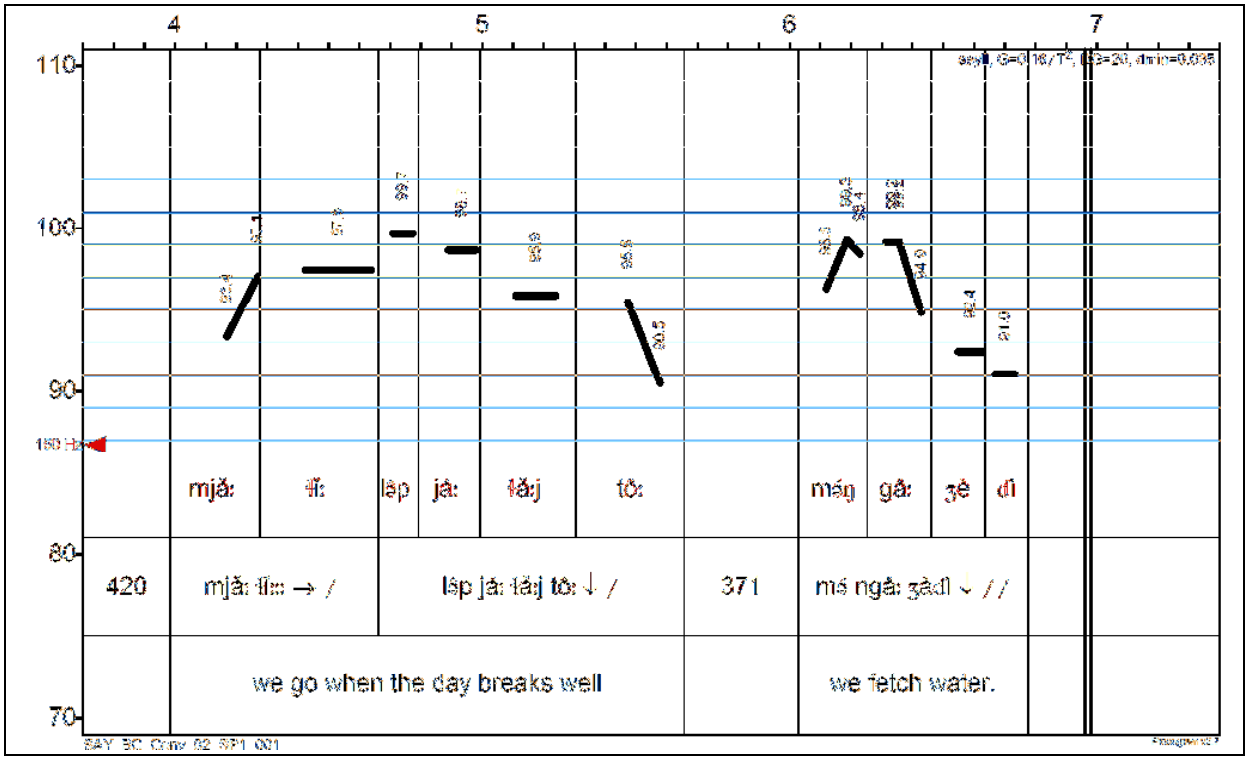
IV. A few examples analysed

Note concerning the representation of intonation: by default : the minor break / is associated with a Level toneme ; →/ is simplified to /. Likewise, the major beak "//" is associated with a Fall ↓; ↓// is simplified to //, and ↓// will be used to mark the end of a period. However, the transcription inside the diagrams has kept the full notation with its redundancies.

A. Complex paratone. /___↑/ ___/ ___/ ___↑↓/ ___/ ___↓/ ___//

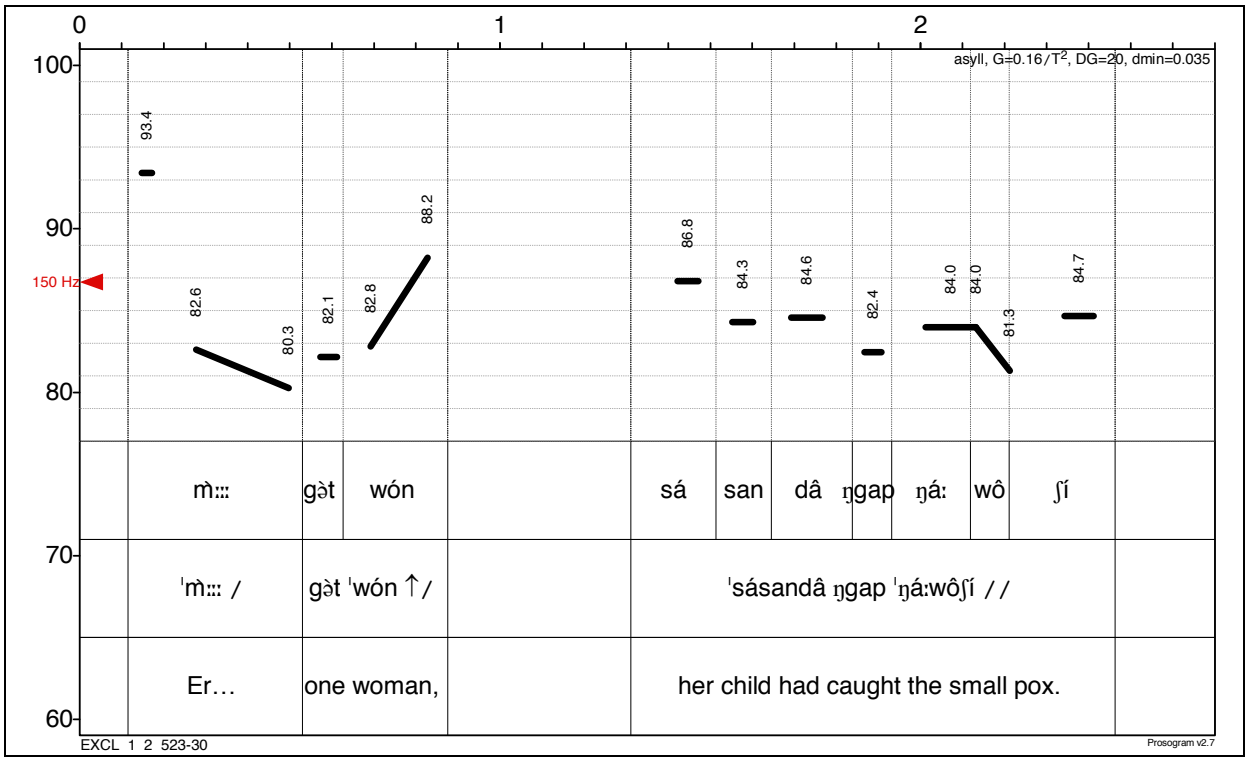
Example 13. SAY_BC_Conv_02_SP1_001

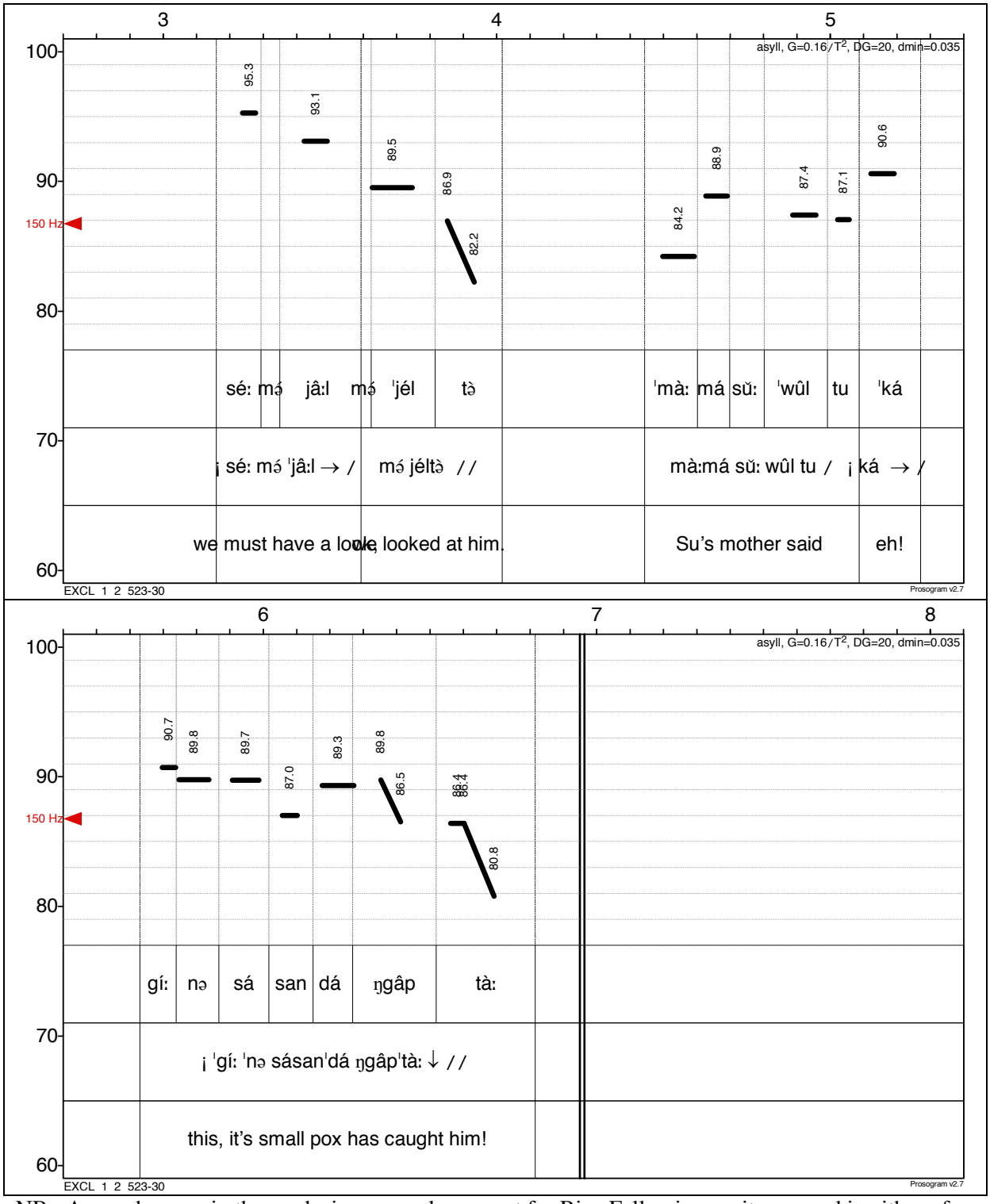




B. Period. ___/ ___↑/ ___/ ¡___/ ___// ___/ ¡___/ ¡___↓//

Example 14. Haruna, excl_1_2_523-30





NB : As can be seen in the predecing examples except for Rise-Fall, minor units can end in either of the other intonemes : Level, Fall and Rise. These are not phonologically determined ; they do not depend on the lexico-grammatical tone of the final syllable. Moreover, in the CorpAfroAs corpus, the Prosogram representations show pitch movements inside intonation units (represented by slanted bars), that do not correspond to contour tones. These internal intonemes often (but not always) correlate with the same IU-final intonemes. Futher study is needed to see if their distribution can be associated with a specific informational function.

V. Conclusion

Two conclusions can be drawn of a preliminary exam of the few elements discussed in this paper about Zaar intonation:

If we refer to the typology sketched by T. Bearth (1998:80-1), which distinguishes between two types of languages, i.e. (i) those that stack intonation patterns over lexico-gramatical tones and (ii) those that express intonation at the periphery of the utterance, Zaar would be a mixed language, with both internal intonation (with upstep and downstep inducing pitch-raising over whole intonation units) and peripheral intonation (with Rise and Fall final intonemes). This could be further developed if the final intonemes are confirmed to be correlated by anticipatory Rises and Falls inside the intonation units.

Beyond the variations in the location of intonemes (whether peripheral or contiguous with the whole intonation unit), the same general pragmatic interpretation of intonation contours seems to hold for Zaar, as well as for Toura, and English for that matter: « […] une unité discursive tonalement décalée [downstepped] par rapport aux unités voisines est caractérisée soit comme information d'arrière-plan par rapport à ces dernières soit comme reprise d'un déjà-dit » (Markus 2006 :117, in Bearth *op.cit*.).

VI. References

Bearth, Thomas. "Tonalité, Déclinaison Tonale Et Structuration Du Discours - Un Point De Vue Comparatif." In Les Unités Discursives Dans L'analyse Sémiotique: La Segmentation Du Discours, edited by Gustavo Quiroz, Ioanna Berthoud-Papandropoulou, Evelyne Thommen and Christina Vogel. Frankurt/M., Berlin, Bern, NewYork, Paris, Wien: Peter Lang. 73-87, 1998.

Beckman, Mary E and Jeniffer J. Venditti. Tone and Intonation. In: Hardcastle, William J., Laver, John and Gibbon, Fiona E. (eds.). The Handbook of Phonetic Sciences. 2nd edition. (Blackwell Handbooks in Linguistics.) Chichester: Wiley-Blackwell. 603-650, 2010

Dehé, Nicole & Yordanka Kavalova (eds.). Parentheticals (*Linguistik Aktuell/Linguistics Today* 106). Amsterdam & Philadelphia: John Benjamins, 2007

Dehé, Nicole. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45: 569–615, 2009

Gussenhoven, C. Intonation and interpretation : Phonetics and Phonology.*Speech Prosody2002, An International Conference*, Bel, B. And Marlien, I. (eds), Aix-en-Provence, Laboratoire Parole et Langage,.CNRS and Université de Provence. 47-75, 2002.

Hirst, Daniel. Lexical and non-lexical tone and prosodic typology, In TAL-2004 (International Symposium on Tonal Aspects of Languages With Emphasis on Tone Languages, Beijing, China, March 28-31, 2004), 81-88, 2004. <http://www.isca-speech.org/archive/tal2004/tal4_081.pdf>

Inkelas, Sharon and Leben, William R. 1990. Where phonology and phonetics intersect: the case of Hausa intonation. Between the Grammar and the Physics of Speech, ed. by Mary E. Beckman and John Kingston, 17-34. New York: Cambridge University Press.

Inkelas, Sharon, Leben, William R. and Cobler, Mark. 1987. The phonology of intonation in Hausa. Proceedings of the Annual Meeting of the North Eastern Linguistic Society.327-41.

Izre'el, Shlomo and Mettouchi, Amina. "Representation of Speech in CorpAfroAs. Transcriptional Strategies and Prosodic Units" (this volume).

Lambrecht, Knud. *Information Structure and Sentence Form : Topic, Focus, and the Mental Representations of Discourse Referents*, *Cambridge Studies in Linguistics (71)*. Cambridge England ; New York: Cambridge University Press, 1994.

Leben, William R. 1985. Intonation in Chadic languages. Précis from the Fiteenth Conference on African Linguistics [UCLA, 1984], ed. by Russell G. Schuh, 191-95. Los Angeles: Department of Linguistics, UCLA.

Leben, William R. 1989. Intonation in Chadic: An overview. Current Progress in Chadic Linguistics. Proceedings of the International Symposium on Chadic Linguistics, Boulder, Colorado, 1-2 May, 1987, ed. by Zygmunt Frajzyngier, 199-218. Amsterdam/Philadelphia: John Benjamins.

Lindau, Mona. 1983. Intonation in Hausa. Abstracts of the 10th International Congress of Phonetics Sciences, 611. Dordrecht: Foris.

Lindau, Mona. 1984. Testing a model for Hausa intonation. Working Papers, Lund University, Department of Linguistics.145-63.

Lindau, Mona. 1986. Testing a model of intonation in a tone language. Journal of the Acoustical Society of America.757-64.

Lindsey, Geoffrey A. 1985. Intonation and interrogation: tonal structure and the expression of a pragmatic function in English and other languages, UCLA: PhD.

Markus, M. (2006). English and German Prosody - A Contrastive Comparison. *Prosody and syntax: cross-linguistic perspectives*. Y. Kawaguchi, I. Fónagy and T. Moriguchi. Amsterdam/Philadelphia, John Benjamins: 103-124.

Rialland, Annie and Robert, Stéphane. 2003. Intonation et focalisation en wolof : Economie d'un système. Fonction et moyens d'expression de la focalisation à travers les langues, 161-85 Louvain: Peeters.

Schneeberg, Nan. "Sayanci Phonology." Ph.D., Indiana University, 1974.

Wichmann, Anne. *Intonation in Text and Discourse. Beginnings, Middles and Ends*. Harlow, London, etc. : Longman, 2000.