Phonology and intonation

Caroline Féry, Sam Hellmuth, Frank Kügler, Jörg Mayer,
Ruben Stoel and Ruben van de Vijver

University of Potsdam

The encoding standards for phonology and intonation are designed to facilitate consistent annotation of the phonological and intonational aspects of information structure, in languages across a range of prosodic types. The guidelines are designed with the aim that a non-specialist in phonology can both implement and interpret the resulting annotation.

1 Preliminaries

This paper presents the conventions of the SFB 632 concerning the archiving of phonological and intonational information. Targets of the standardization are:

- to archive language data which are not understandable for every database user offhand in a comprehensive way;
- to enable phonological/intonation queries in the database, e.g., "search for a high pitch accent in a focused position".

Some of the annotation conventions presented below are only comprehensible to phonologists. If you feel uncomfortable with some of the terms ('mora', 'trochee' could be some), ask for advice rather than using them arbitrarily.

2   Levels declaration

The group "Phonology and Intonation" has prepared guidelines for the levels of phonology and intonation. The levels and the corresponding layers are declared in the following table:

Table 1: Levels and layers

Level	Layer	Name
	General information	INFO
Sound	Orthographic transcription	WORDS
	Phonemic transcription (syllables)	PHONES
Metrical Structure	Stress	STRESS
	Accent	ACCENT
	Mora	MORA
	Foot	FT
Prosodic structure	Phonological Words	PW
	Phonological Phrases	PHP
	Intonational Phrases	IP
	Underlying tones	TONES
Tones and Intonation	Surface tones	SURFACE
	Phonetic tones	PHONTONES

2.1   Related existing standards

Following standards have been considered for the development of the current guidelines.

Table 2: Existing standards

	ToBI	IViE	GAT
graphemic transcription	x	x	(x)
phonetic transcription	(c)	(x)	(x)
moraic layer	-	-	-
syllabic layer	-	-	-
accent domain / phonological phrase	-	-	-
intermediate layer	x	-	-
intonation phrase layer	x	x	(x)
utterance layer	(x)	(x)	(x)
underlying tonal layer	-	-	-
surface tonal layer	x	x	x
phonetic layer	-	x	-
prominence layer	-	x	(x)

There are significant differences between the first two and the last of the three standards: ToBI & IViE were developed for tonal transcription of intonation; GAT was developed for transcription of conversational speech data. Although the GAT system concerns a different purpose, it also contains a section about intonational transcription. For the purposes of the SFB 632, layers of prosodic information have to be included in the annotation system.

2.1.1   ToBI

ToBI stands for "Tone and Break Indices". This standard is based on the theoretical groundwork of intonational phonology by Pierrehumbert (1980), and provides conventions for transcribing the intonation and prosodic structure of

spoken utterances in a language variety. Yet, a ToBI system represents an already conducted phonological analysis of the intonation of a certain language.

For more information of the general idea of ToBI see:

http://www.ling.ohio-state.edu/~tobi/

For labeling conventions of English ToBI (Beckman & Ayers 1997) see:

http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/

2.1.2   IViE

IViE stands for "Intonational Variation in English". This standard was developed on the basis of ToBI. The purpose is to ensure comparable transcription of several varieties of English using a single labeling system. The system has been developed within the research project "English Intonation in the British Isles" (Grabe & Nolan, see: http://www.phon.ox.ac.uk/~esther/ivyweb/). In addition to ToBI, IViE allows for transcription of prominence and phonetic variation. For more information on the labeling conventions see:

http://www.phon.ox.ac.uk/~esther/ivyweb/guide.html

2.1.3   GAT

GAT stands for "Gesprächsanalytisches Transkriptionssystem" (Selting et al. 1998). Its purpose is to provide a standard for transcription of conversational speech. The sequential structure of a conversation is iconically represented in the transcription, where each speaker's turn starts with a new transcription line. It is not clear whether a speaker's turn coincides with a prosodic unit, i.e. whether a one-to-one correspondence exists between a turn and an intonation phrase for instance. In addition to the segmental graphical level, GAT provides labels for prosodic annotation, which are however very impressionistic and rough.

For more information see:

http://www.fbls.uni-hannover.de/sdls/schlobi/schrift/GAT/ and

http://www.fbls.uni-hannover.de/sdls/schlobi/schrift/GAT/gat.pdf

3   Level I: Sound + Transcription

3.1   Declaration

The level consists of 3 layers: Orthographic transcription (WORDS), phonemic transcription (PHONES), and general information (INFO). Words can be annotated in transliteration or in the original orthography. The obligatory layer of broad phonemic transcription (PHONES) is done in either IPA or SAMPA. If you decide to use SAMPA and no version of SAMPA for your language is available, use X-SAMPA (for more information about SAMPA and X-SAMPA, see: http://www.phon.ucl.ac.uk/home/sampa/home.htm). Specify which phonetic alphabet you are using in the general information layer (INFO). You should also specify the sound file name in the INFO layer. Other optional entries to this layer are recording conditions, speaker characteristics, etc.

3.2   Objects of annotation

The annotated sentences must be available as sound files in formats that are readable by PRAAT, preferably in the .wav format. Other possible formats are: aiff, aifc (but not compressed aifc files), au files (NeXT/Sun) and NIST files: big-endian, little-endian, μ-law, A-law, Polyphone (NIST files compressed with **shorten** are not supported).

3.3   Tagset declaration

The WORDS and INFO layers are free text fields which require no tagset. The PHONES layer tagset comprises all IPA or (X-)SAMPA symbols, plus one tag

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)
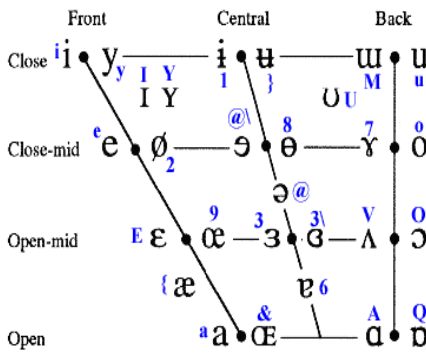
CONSONANTS (PULMONIC) — WITH X-SAMPA EQUIVALENTS IN BLUE

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
ʘ Bilabial	ɓ Bilabial b_<	' as in:
ǀ Dental	ɗ Dental/alveolar d_<	p' Bilabial p_>
ǃ (Post)alveolar	ʄ Palatal J_<	t' Dental/alveolar t_>
ǂ Palatoalveolar	ɠ Velar g_<	k' Velar k_>
ǁ Alveolar lateral	ʛ Uvular G_<	s' Alveolar fricative s_>

SUPRASEGMENTALS

%foUn@"tIS@n

ˈ	Primary stress	ˌfoʊnəˈtɪʃən
ˌ	Secondary stress	
ː	Long	eː e:
ˑ	Half-long	eˑ e:\
̆	Extra-short	ĕ e_X
.	Syllable break	ɹi.ækt r\i.{kt
|	Minor (foot) group	
‖	Major (intonation) group	
‿	Linking (absence of a break)	

TONES & WORD ACCENTS

LEVEL		CONTOUR	
e̋ or ˥	Extra high _T	ě or ˇ	Rising _R
é ˦	High _H	ê	Falling _F
ē ˧	Mid _M	᷇	High rising _H_T
è ˨	Low _L	᷄	Low rising _B_L
ȅ ˩	Extra low _B	᷈	Rising-falling _R_F
↓ Downstep !		↗ Global rise <R>	
↑ Upstep ^		↘ Global fall <F>	

VOWELS

	Front	Central	Back
Close	i y	ɨ ʉ	ɯ u
	I Y	1 }	M u
	ɪ ʏ		ʊ U
Close-mid	e ø	ɘ ɵ	ɤ o
	2	@\ 8	7 o
		ə @	
Open-mid	ɛ œ	ɜ ɞ	ʌ ɔ
	E 9	3 3\	V O
	æ	ɐ	
	{	& 6	
Open	a ɶ		ɑ ɒ
	a OE		A Q

Where symbols appear in pairs, the one to the right represents a rounded vowel.

OTHER SYMBOLS

- ʍ Voiceless labial-velar fricative — W M
- w Voiced labial-velar approximant — w W
- ɥ Voiced labial-palatal approximant — H ɥ
- ʜ Voiceless epiglottal fricative — H\ H
- ʢ Voiced epiglottal fricative — <\ ʢ
- ʡ Epiglottal plosive — >\ ʡ
- ɕ ʑ Alveolo-palatal fricatives — s\ z\
- ɺ Alveolar lateral flap — J ɺ
- ɧ Simultaneous ʃ and x — x\
- Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.
- k͡p t͡s — k_p t_s

DIACRITICS

X-SAMPA diacritics come after symbols, e.g. n_0
Diacritics may be placed above a symbol with a descender, e.g. ŋ̊

̥	Voiceless	n̥ d̥	̤	Breathy voiced	b̤ a̤	̪ Dental t̪ d̪
̬	Voiced	s̬ t̬	̰	Creaky voiced	b̰ a̰	̺ Apical t̺ d̺
ʰ	Aspirated	tʰ dʰ	̼	Linguolabial	t̼ d̼	̻ Laminal t̻ d̻
̹	More rounded	ɔ̹	ʷ	Labialized	tʷ dʷ	̃ Nasalized ẽ
̜	Less rounded	ɔ̜	ʲ	Palatalized	tʲ dʲ	ⁿ Nasal release dⁿ
̟	Advanced	u̟	ˠ	Velarized	tˠ dˠ	ˡ Lateral release dˡ
̠	Retracted	i̠	ˤ	Pharyngealized	tˤ dˤ	̚ No audible release d̚
̈	Centralized	ë	̴	Velarized or pharyngealized	ɫ (or velarized l: 5)
̽	Mid-centralized	e̽	̝	Raised	e̝ (ɹ̝ = voiced alveolar fricative)
̩	Syllabic	ɹ̩	̞	Lowered	e̞ (β̞ = voiced bilabial approximant)
̯	Non-syllabic	e̯	̘	Advanced Tongue Root	e̘
˞	Rhoticity	@` ɚ	̙	Retracted Tongue Root	e̙

for pauses: <P> (you can use the <P>-tag also in the WORDS layer). The following table has all the phonetic symbols in IPA and in X-SAMPA. The table can be downloaded from the following site:

http://www.diku.dk/hjemmesider/studerende/thorinn/xsamchart.gif

In order to use phonetic symbols in PRAAT one has to have SIL Doulos IPA 1993 font installed. It can be downloaded from the PRAAT site (http://www.praat.org). A tutorial on the use of phonetic symbols in PRAAT can be found in the manual of the PRAAT program itself or here:

http://www.fon.hum.uva.nl/praat/manual/Phonetic_symbols.html.

Some illustrative examples are given below. Note the mismatch between word and syllable boundaries:

(1)   Dutch using IPA

<INFO>	Sound file name, IPA								
<WORDS>	Marie	geeft	Kees	een	klap	op	zijn	kop	.
<PHONES>	maː	riˑ	ɣeːft	keːs	ən	klɑp	ɔp	zɛin	kɔp	
<TRANS>	Marie hits Kees on the head.								

(2)   Dutch using SAMPA

<INFO>	Sound file name, SAMPA								
<WORDS>	Marie	geeft	Kees	een	klap	op	zijn	kop	.
<PHONES>	ma:	ri.	Ge:ft	ke:s	@n	klAp	Op	zEin	kOp	

(3)   French

<INFO>	SAMPA						
<WORDS>	La	maison	s'	est	écroulée		.
<PHONES>	la	mE	zo~	sE		te	kru	le
<TRANS>	The house collapsed.						

3.4 Instructions

INFO

- Specify the sound file name and the phonetic alphabet used in the

PHONES layer (IPA/SAMPA).

- If necessary add additional information.

WORDS

- Determine word boundaries and provide an orthographic transcription in the relevant interval.

- You can annotate significant pauses with the <P>-tag.

PHONES

- The transcription should be broad (phonemic). For example, in English there is no phonemic contrast between aspirated and non-aspirated stops, so it is not necessary to mark aspiration in the transcription.

- Obtain syllable boundaries and enter the appropriate phonetic symbols for each syllable either in IPA or in SAMPA. Though it is usually quite easy to decide on the number of syllables, the decision as to the exact location of their boundaries may be trickier. If there is consensus about syllabification, use the standards. Otherwise, use your own intuitions.

- Do not mark stress at this level; this should be done at the 'STRESS' level (see section 3.3).

4   Level II: Metrical structure

4.1   Objects of annotation

Lexical stresses are important to locate because they carry the pitch accents related to sentence accent. We use the term 'stress' to refer to the prominent syllable in a word, disregarding whether the syllable is realized as prominent or not, and the term 'accent' to refer to realization of stresses by means of a pitch accent. The phonetic and phonological properties of individual pitch accents are the object of section 5. Here, we deal only with the abstract position of lexical

stress and whether or not potential stresses are realized by means of pitch accents. Note that an accent may sometimes be realized on a syllable which does not carry lexical stress: e.g. 'I said stalagMITE, not stalagTITE'. Some languages have no lexical stress, such as many tone languages, and investigators of these languages can omit the STRESS level; however there are languages which do not have lexical stress but which do have accents, such as French, in which case the ACCENT layer only should be used. Do not transcribe lexical tones at this level (see instead the Tones and Intonation level).

For languages with lexical stress, the STRESS layer should be 'abstract'. It should consider words spoken in isolation, and not the phonetic realization in context, since this latter realization is subject to accent shifts, accent deletion and the like.

4.2 Tagset declaration

Table 3: Tagset declaration for metrical structure

tag	meaning	short description
1	Primary stress	Most prominent syllable of the language
2	Secondary stress	Second most prominent syllable of the language
3	Tertiary stress	Third most prominent syllable of the language

4.3 Instructions

STRESS

- If the object language has lexical stress, then annotate the most prominent syllables (or mora) of the language with 1.

- If the object language has lexical stress, then annotate the second most prominent syllables (or mora) of the language with 2.

- Usually two layers of lexical stress are sufficient. In general, tertiary stress is not necessary. The third most prominent syllables (or mora) of the language can be annotated with 3 if you feel that it is important. All lower layers should be annotated with 3 as well.

- Unstressed syllables are not especially annotated.

- It is not necessary to annotate mono-syllabic function words.

ACCENT

- If the object language has lexical stress, annotate the stresses realized by means of a pitch accent with 1.

- If the object language does not have lexical stress but does have accents, annotate syllables realized with a pitch accent with 1.

- Unaccented syllables are not especially annotated.

(4)   German, SAMPA

<WORDS>	Lena		verfolgt		den	Mann	mit	dem	Motorrad		
<PHONES>	le:	na	v6	fOlgt	d@n	man	mIt	d@m	mo	to:	Rat
<STRESS>	1	2		1		1				1	2
<ACCENT>	1					1				1	

5   Level III: Prosodic structure

5.1   Declaration

For the sake of information structure, it is crucial to annotate the higher levels of prosody which are relevant for information structure. In most cases, these are the Phonological Phrase (PHP) and the Intonation Phrase (IP). It is also obligatory to specify syllable boundaries (SYL). The other layers (Phonological Word, Foot, and Mora) are required only if the annotators are phonologically confident. Constituents larger than a sentence (intonation phrase) are not considered here.

5.2 Related standards

Terminological conventions differ in some respects. Phonological Words are sometimes called 'Prosodic Words'. Phonological Phrases can be called 'Accentual Phrases'. 'Clitics Group' is sometimes considered to be a necessary layer, but it is not part of the required annotations.

5.3 Tagset declaration

Table 4: Tagset declaration for metrical structure

Layer	Tags	Meaning	Short Description
IP	IP	Intonation Phrase	A phonological constituent roughly the size of a sentence
PHP	PP	Phonological Phrase	A phonological constituent roughly the size of a maximal projection
	aPP	Abstract PP	See below for a description
	rPP	Realized PP	
PW	PW	Prosodic Word	A phonological constituent roughly the size of a grammatical word
FT	F	Foot	A metric unit ideally consisting of two syllables, one of which is strong and the other weak (trochee or iamb)
MORA	M	Mora	A weight unit of the syllable (its nucleus and possibly the next consonant)

(5)   English (adapted from Gussenhoven (2002: 271))

<WORDS>	Too	many		cooks	spoil	the	broth	.
<PHONES>	tu:	mE	ni	kUks	spOIl	D@	brOT	
<FT>	F	F		F	F		F	
<PW>	PW	PW		PW	PW	PW		
<PHP>	PP				PP	PP		
<IP>	IP							

(Note that, in this analysis, the word 'the' is not part of any foot.)

5.4   Instructions

5.4.1   Mora (MORA)

- Only some languages use the mora actively in the phonology and distinguish tones or accent placement in terms of moras. An example is Japanese in which the default high tone in unaccented words is on the second mora, which can be part of the first or the second syllable.

H   H

| |

o͟osutária 'Austria'   gar͟asudama 'glass beads'

μμ   σ σ

Syllable:

- Syllable boundaries should be annotated at the PHONES layer (see section 2.4).

Foot (FT):

- According to most researchers, trochees (strong - weak) and iambs (weak - strong) are the only universal metrical feet. Trochees may be based on moras or on syllables. This gives the following universal foot inventory (Hayes 1995):

     a. Syllabic trochee: (canonical form: 'σσ)

     b. Moraic trochee: (canonical form: 'μμ)

     c. Iamb: (canonical form: σ'σ)

- If the metrical footing is described in the literature, use this standard, otherwise use your own.

- Feet are used for identification of lexical stress. It is often sufficient to locate primary (and secondary and tertiary) stress without any foot structure. If you do not feel confident, do not annotate this layer.

Phonological Word (PW):

- Prosodic Word is a constituent which is more or less isomorphic to the grammatical word, except for cliticization. In languages with lexical stress, one syllable has main prominence. PW are often domains for syllabification and can stand alone.

- The identification of this domain is not crucial for information structure. Use it only if you feel confident about it, or if it is important for the higher layers.

Phonological phrase (PHP):

- A phonological phrase is a domain for syntactic accents and for phrasing. Depending on the presence of accents, a PHP contains a main accent or an element which is more prominent than the other ones. A PHP is often delimited by boundaries, which can be tonal, segmental or durational.

- If you have very good knowledge or intuitions about this layer, you may distinguish between an abstract PHP and a realized PHP. An abstract PHP is defined by the syntactic structure.

Intonation phrase (IP):

- An intonation phrase is the domain of the tonal realization. In languages with lexical stress and/or sentence accent, it is the domain of the main

accent. In most cases, it follows the syntactic structure closely, and an IP is often isomorphic to a sentence. Embedded clauses usually form separate IPs, as do parenthetical clauses. A vocative or a tag (like 'isn't it') also usually form separate phrases. Lists, elliptical constructions, inverted discontinuous constituents, may be in separate IPs as well as cleft constructions, topicalizations and extrapositions.

- In intonation and pitch accent languages always, and in tone languages most of the time, intonation phrases are delimited by a boundary tone, meaning that a word has a different intonation (falling or rising) when it is phrase-final than when it is phrase-medial.

6   Level IV: Tone and Intonation

6.1   Introduction

The intonation of a language can be transcribed using the ToBI framework. ToBI systems are available for English, German, Dutch, Greek, Japanese, Korean etc. See Jun (2005) for an overview of ToBi in different languages. We propose three layers of tonal transcription: Two of these layers (TONES and SURFACE) capture phonological descriptions of tones. If a ToBI system exists for the language, please use it. Otherwise follow the directions in the remainder of this document. See also Ladd (1996) and Gussenhoven (2004) for descriptions of intonation from a theoretical perspective. The third layer (PHONTONES) concerns a phonetic description in the style of the IviE system (see below).

6.2   Declaration

In tone languages, lexical items are associated with tones. The underlying tone of each syllable or word should be represented in the TONES layer. If words

associated with tones occur in combination, certain tones may change on the surface due to phonological process or because of constraints like the OCP (cf. 5.2.1.). In intonation languages, tonal processes such as tonal linking may also change the underlying tonal representation of pitch accents (cf. 5.2.2.). The output of these and similar processes is annotated in the SURFACE layer. If very little is known about a language one might at first describe surface pitch patterns around prominent syllables (in intonation languages) or tone-bearing units (in tone languages) in the PHONTONES layer. The phonetic layer may thus help capture surface variation which leads into phonological abstraction in further analysis.

6.2.1   Underlying tones (TONES)

The layer TONES comprises lexical tones (as in Chinese), lexical accents (as in Japanese or Swedish), intonational tones (as in English or German), and boundary tones. Tone bearing units (TBUs) can be syllables (most languages) or moras (Japanese for instance). Tonal layers should be indicated according to the following standards:

- use the labels H (high) and L (low); possibly also M (mid)
- alternatively, use tonal labels according to existing conventions, such as tone 1, 2, 3, 4 in Mandarin Chinese.

Please use the following conventions (see also ToBI conventions):

Table 5: Tagset declaration for metrical structure

H* / L*	high-tone / low-tone pitch accent
L*H	rising pitch accent
H*L	falling pitch accent
L*HL	rise-fall pitch accent (other combinations possible)
!H*	downstepped high-tone pitch accent

H- / L-	high/low boundary tone associated with Phonological Phrase
H% / L%	high/low boundary tone associated with Intonational Phrase right edge
%H / %L	high/low boundary tone associated with Intonational Phrase left edge
LH%	rising boundary tone (other combinations possible)
*?	uncertain if this is a pitch accent
X*	this is a pitch accent but I don't know which one
?- / ?%	uncertain if this is a boundary tone
X- / X%	this is a boundary tone but I don't know which one

Notes:

- A pitch accent is a tone that is associated with the stressed syllable of a word. A boundary tone marks the beginning or end of a prosodic domain, such as a Phonological Phrase or an Intonational Phrase.

- The transcription of both types of tones should be phonological rather than phonetic, thus do not include more detail than needed to make categorical distinctions.

(6)   English (from Gussenhoven 2002: 271)

<WORDS>	Too	many		cooks	spoil	the	broth	.
<PHONES>	tu:	mE	ni	kUks	spOIl	D@	brOT	
<PHP>	PP				PP	PP		
<IP>	IP				IP			
<TONES>	H*			L*H	H*		H*L	L%

6.2.2   Surface tones (SURFACE)

The same conventions apply on the SURFACE layer as for the TONES layer (see above). The significant difference between the two layers is that the SURFACE layer gives information about tones that have undergone phonological adjustments, such as tone sandhi and tonal linking. If in doubt, assume that these two layers are identical.

(7)   Mandarin (The underlying tone 1 of *yi* surfaces as tone 4, and the underlying tone 3 of *suo* surfaces as tone 2; the other tones have the same underlying and surface form)

<ORTHOGR>	一本书惹怒了所有的人.									
<TRANS>	A book annoyed everyone.									
<GLOSS>	one	CL	book	annoy		ASP	everyone			
<WORDS>	yi	ben	shu	re-nü		le	suo-you-de-ren			.
<PHONES>	i	pən	ʂu	rə	ny	lə	suə	iəu	tə	rən	.
<TONES>	1	3	1	3	4	0	3	3	0	2
<SURFACE>	4						2			

6.2.3   Phonetic description (PHONTONES)

The PHONTONES layer provides space to label phonetic surface variation. This optional layer allows for a first analysis-free inspection of the data. If no intonational information of a language is available, this layer may be used to build up generalisations of surface tonal patterns. The systematic comparison of utterances in the same context provides insights in the phonological structure of a language. An additional effect is the outcome of phonology-phonetic mapping rules, and how they differ between languages (see IViE for the use of this layer). This layer comprises the pitch contour around prominent syllables, (Implementation domain: preaccented, accented plus following syllables up to the next accented one).

Table 6: Available labels from IViE

H / h	high pitch	⎤	Upper case is used for the accented syllable, lower case for preceding and following unaccented syllables.
L / l	low pitch	⎬	
M /m	mid pitch (optional)	⎦	
i	end of an implementation domain		
-	interpolation		

An implementation domain may contain maximally four labels (preaccentual, accented, postaccentual and final domain syllable), minimally two (accented and pre- *or* postaccentual syllable).

Figure 1: Illustrative examples of an implementation domain

(from http://www.phon.ox.ac.uk/~esther/ivyweb/guide.html)
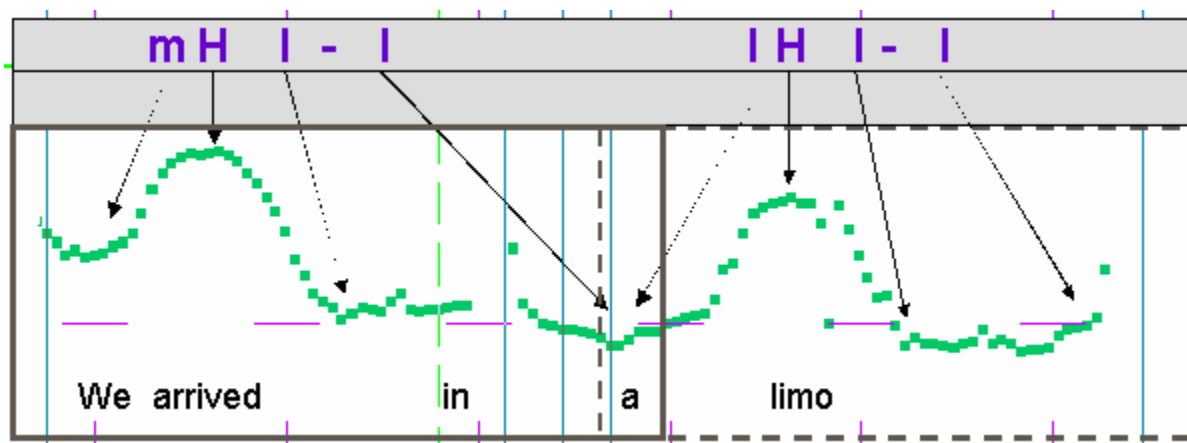
7   Guidelines for phonological analysis

7.1   Checklist

To conduct a phonological analysis, you should try to answer the following questions:

- Does the language have lexical tones or pitch accents? A lexical tone is a specific melody (like a high tone, a low tone, a falling tone) associated with syllables or moras of words and contributing to the lexical meaning of the word. A pitch accent is found in some languages (for instance Japanese and Swedish). It has the same function as a lexical tone, but pitch-accent languages typically have only one pitch accent, whereas tone languages typically have several tones.

- What kind of tonal entities exist in the language? (Section 7.2)

- Are focus and topic expressed by intonation (i.e. an accented syllable)?

(Section 7.3)

- Is a narrow focus accent possible, even on an unstressed syllable (an English example: "I didn't say coLA I said coDA")? (Section 7.3)

- How is a given or 'backgrounded' element realized in the language? (Section 7.4)

- What is the relation between sentence type and intonation? In particular, can you say that a question is generally expressed with a rising intonation, and a declarative sentence with a falling intonation? (Section 7.5)

7.2 Tonal entities

7.2.1 Lexical tones and lexical pitch accents

Lexical tones and lexical pitch accents change the meaning of a word, and they are thus specified in the lexicon, contrary to intonational tones. A language has lexical tone if most syllables have a lexical tone associated with them. Examples include Chinese and many African and Amerindian languages (see the Chinese example above). A language has lexical pitch accent if even one syllable of a word (or phrase) has a tone associated with it. Examples include Japanese and Swedish.

Example Swedish:

(8a)   Accent 1 word (HL*), from Bruce (1977), tonal labels adapted

<WORDS>	Man	vill	anamma		nåra	längre		[1]nummer	
<PHONES>	$	$	$	$	$	$	$	$	$	$
<TONES>			HL*			HL*		HL* H-	L%
<STRESS>			2			2		1	
<TRANS>	"One wants to receive some longer numbers"								

(8b)   Accent 2 word (H*L), from Bruce (1977), tonal labels adapted

<WORDS>	Man	vill	lämna		nåra	långa		²nunnor	
<PHONES>	$	$	$	$	$	$	$	$	$
<TONES>			H*L			H*L		H*L	H- L%
<STRESS>			2			2		1	
<TRANS>	"One wants to leave some long nuns."								

Example: Mandarin Chinese

- lexical tone (1, 2, 3, 4); A single word (like *ma*) has completely different meanings according to the lexical tone associated with it.

- no pitch accent

- L and H boundary tones

- raised, expanded and compressed pitch range

Example: Japanese

- lexical pitch accent

- only one type of pitch accent, H*+L

- L and H boundary tones associated with the Accentual Phrase (AP)

- H, LH, and HL boundary tones associated with Intonational Phrase (IP)

          H*L   H*(L)

           | | |

      a. hána 'a name'   b. haná 'flower   c. hana 'nose

7.2.2   Intonational tones

Intonational tones are assigned at a 'post-lexical' level. Current theories of intonational phonology distinguish two types of intonational tones: (post-lexical) pitch accents and boundary tones. Both of these may be analyzed as sequences

of simple tones (H and L). A pitch accent is associated with the stressed syllable of a word, and a boundary tone with the beginning or end of a prosodic domain. Not all languages have pitch accents, but most languages appear to have boundary tones. Some (tone) languages use pitch range features instead of pitch accents.

There are typically many syllables that do not have a tone associated with them, and the pitch between two tones is filled in by interpolation, e.g. between a high tone and a low tone there will be a gradual decline in pitch, which does not need to be specified in the phonological transcription.

Example: German

- pitch accents[16]: L*, H*, LH*, L*H, H!H*, and H*L (or HL*)

- boundary tones: L and H

(9)   (Example adapted from Féry 1993)

<WORDS>	Lena		verfolgt		den	Mann	mit	dem	Motorrad		
<PHONES>	le:	na	v6	fOlgt	d@n	man	mIt	d@m	mo	to:	Rat
<IP>	IP										
<TONES>						H*L				H*L	
<SURFACE>						H*				H*L	

7.3 Information structure

To study the relation between intonation and information structural categories such as topic or focus you might take a sentence and have it read with several different focus structures.

7.3.1 Focus structure

To establish whether intonation plays a role for focus assignment in a language, try sentences such as the following:

---

[16] A star refers to a pitch accent, and !H is downstep, i.e. the H tone is realized at a lower pitch than the preceding H tone.

a. Why is he so still? [The child is sleeping]_{FOCUS}

b. Who is sleeping? [The child]_{FOCUS} is sleeping/It is the child….

c. What does the child do? The child [is sleeping]_{FOCUS}

d. What is happening with the mother and the child right now?/ What are the mother and child doing right now?

[The child]_{TOPIC} [is sleeping]_{FOCUS} but…

First check whether intonation is involved at all, or whether the language you are investigating only uses morphological markers or syntax to realize different focus structures. If intonation is involved, check if there is any difference between the intonation of these sentences. Is focus marked by an accent or phrasing? Try to answer the following questions:

- Does the topic in (d) get an accent? Is this accent different from the focus accent?

- Does the topic form a prosodic domain of its own (e.g. PHP or IP)?

- Does the language allow non-final focus as in (b)?

- What intonation pattern does the all-new sentence in (a) have?

See section 7.4 to answer these questions.

7.3.2   Narrow and contrastive focus

You should also check if narrow focus is marked by intonation. That is, is it possible to have a pitch accent (or expanded pitch in a tone language) on a non-final word in an XP?

In English, narrow focus can be marked by a pitch accent in all cases:

(10) (a)   Question: Who loves Mary?

Answer: JOHN loves Mary.

(b)   Question: How many apples do you need for your cake?

Answer: FOUR apples.

In certain cases, an accent signaling narrow focus is ambiguous with an accent signaling wide focus. This happens when the word which is narrowly focused is by accident identical to the word bearing default accent in a wide-focused sentences:

(11)   Question: Who does Mary love?

       Answer: Mary loves JOHN.

Narrow focus cannot always be realized intonationally, because in some languages, pitch accent is restricted to specific positions. In Manado Malay, for instance, an XP must have a final accent. For this reason, if focus is non-final in its phrase, it is not possible to mark it with a pitch accent. Compare the following sentences in English and in Manado Malay (accents are marked in upper case):

(12)   (a)   English:

              How many kilos? [THREE]$_{focus}$ kilos.

       (b)   Manado Malay

              Brapa kilo? [tiga]$_{focus}$ KILO.

       It is important to distinguish narrow and contrastive focus. In a contrastive focus, another element is explicitly denied or contrasted.

(13)   (a)   Contrastive focus:

              I don't want a banana, I want an apple.

       (b)   Question: Is it John that Mary loves?

              Answer: No, Mary loves BILL.

(14) Contrast

       A RED flower is more attractive for birds than a WHITE one.

7.4   Deaccentuation, givenness, backgrounding

Sentences also contain parts which are not prominent. This happens because they are repeated from the sentence before, or simply because they are in the neighborhood of a focused constituent. In the following dialogue, the first part of the answer is backgrounded because it just repeats, albeit with different words, something which has been asked in the preceding question.

(15)   Question: What do you want for dinner?

         Answer: [I would like to eat] background [FISH & CHIPS] focus

In the following example, 'children' may be new, and would thus deserve an accent, but because of the proximity of a contrastive accent, it is prone to deaccentuation.

(16)   John has THREE and Bill has FOUR children.

7.5   Sentence type

The distinction between statement, yes-no question, and wh-question is typically expressed by intonation. In some languages, a statement and a yes-no question are distinguished only by intonation. If this is not the case, then try to make sentences of these types that are as similar as possible. In both cases see what the differences in intonation are. Typical differences include:

- the type of final boundary tone (statements often have L, and yes-no questions H);
- a different kind of pitch accent;
- the position of the pitch accent (e.g. does a wh-word get an accent?);
- the overall pitch level may be higher.

7.6   Practical considerations

If possible, the target words in your sentences (i.e. the words that you expect to get an accent or boundary tone) should not be too short. It is best if they have:

- nasals or voiced obstruents rather than voiceless sounds, since there is no pitch without voicing;

- penultimate or antepenultimate stress, so that no pitch accent and final boundary tone will occur in the same syllable.

8   References

Bruce, Gösta. 1977. *Swedish Word Accents in Sentence Perspective*. Lund: Gleerups.

Grabe, Ersther and Nolan, Francis. 1997. *Intonational Variation in the British Isles*. <http://www.phon.ox.ac.uk/~esther/ivyweb/>

Gussenhoven, C. 2002. Phonology of intonation. *Glot International* 6: 271–284.

Gussenhoven, C. 2004. *The phonology of tone and intonation*. Cambridge University Press.

Jun, S.A. (ed.) 2005. *Prosodic typology: The phonology of intonation and phrasing*. Oxford University Press.

Ladd, D.R. 1996. *Intonational phonology*. Cambridge University Press.

Selting, M. et al. 1998. Gesprächsanalytisches Transkriptionssystem. *Linguistische Berichte* 173, S. 91-122.