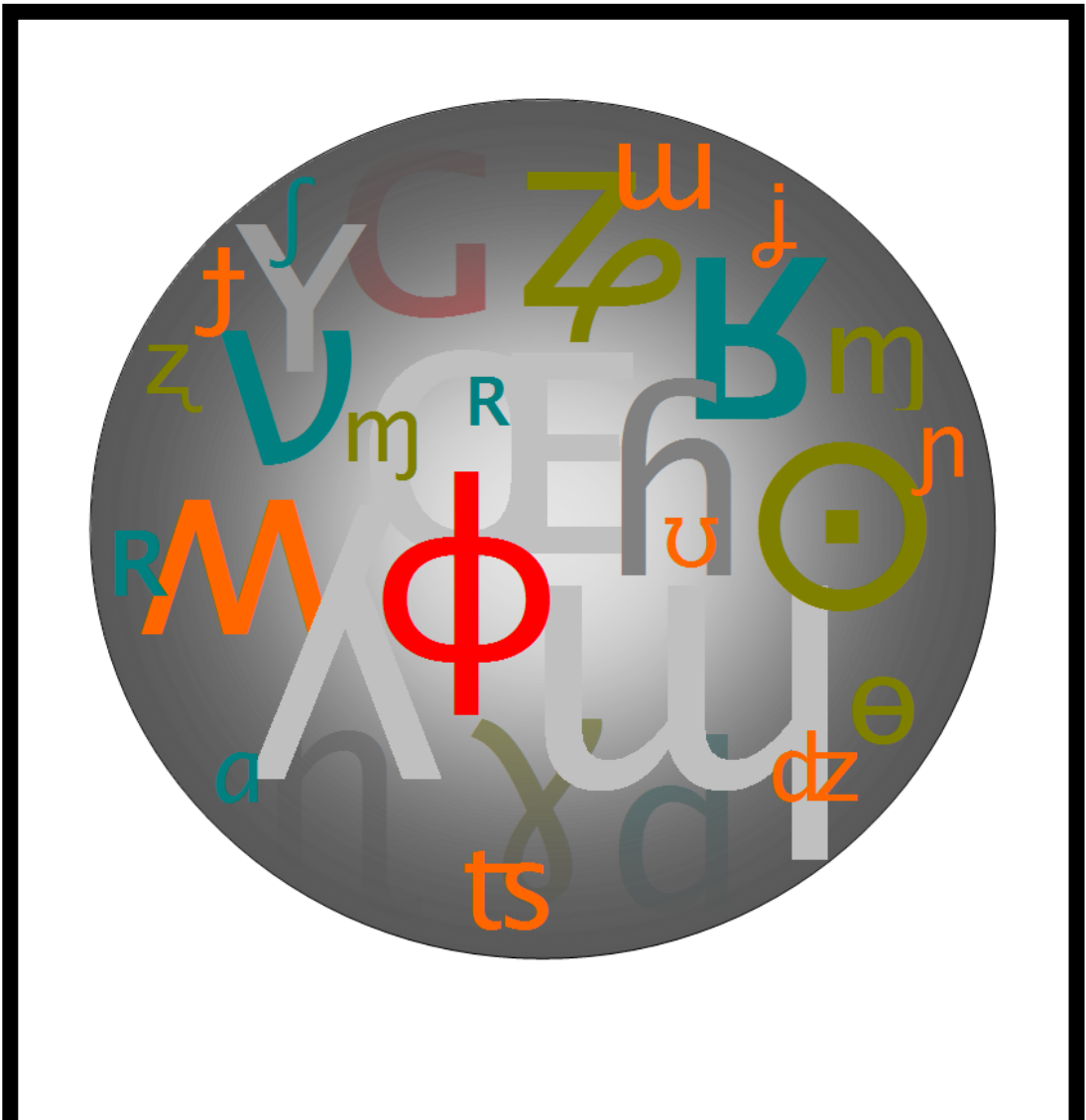


the

Phonetician

A Publication of ISPhS/International Society of Phonetic Sciences



Number 96

2007-II

ISPhS
International Society of Phonetic Sciences

President: Ruth Huntley Bahr

Secretary General:
Mária Gósy

Honorary President:
Harry Hollien

Executive Vice President: Eric Keller

Vice Presidents:

Angelika Braun
Marie Dohalská-Zichová
Mária Gósy
Damir Horga
Eric Keller
Heinrich Kelz
Stephen Lambacher
Asher Laufer
Judith Rosenhouse

Past Presidents:

Jens-Peter Köster
Harry Hollien
William A. Sakow †
Martin Kloster-Jensen
Milan Romportl †
Bertil Malmberg †
Eberhard Zwirner †
Daniel Jones †

Honorary Vice Presidents:

A. Abramson	P. Janota †	P. Ladefoged †	R. K. Potapova	F. Weingartner
S. Agrawal	W. Jassem	A. Marchal	M. Rossi	R. Weiss
L. Bondarko	M. Kloster-Jensen	H. Morioka	M. Shirt	
E. Emerit	M. Kohno	R. Nasr	E. Stock	
G. Fant	E.-M. Krech	T. Nikolayeva	M. Tatham	

Regional Secretaries:

M. Bakalla	Mid-East
M. Dohalská-Zichová	Czech and Slovak Section
C. Paboudjian	Southern Europe
R. K. Potapova	Russia
P. French	United Kingdom
Y. H. Wang	South East Pacific
H. Yamazawa	Japan

Auditor:

Angelika Braun

Treasurer:

Ruth Huntley Bahr

Affiliated Members (Associations):

American Association of Phonetic Sciences	J. Hoit & W. S. Brown
Dutch Society of Phonetics	B. Schouten
International Association for Forensic Phonetics and Acoustics	A. Braun
Phonetic Society of Japan	I. Oshima & K. Maekawa
Polish Phonetics Association	G. Demenko

Affiliated Members (Institutes and Companies):

Kay Elemetrics, Lincoln Park, NJ, USA	J. Crump
Inst. for Advanced Study of the Communication Processes, Univ. Florida, USA	H. Hollien
Dept. of Phonetics, University of Trier, Germany	J.-P. Köster
Dept. of Phonetics, University of Helsinki, Finland	A. Iivonen
Dept. of Phonetics, University of Zürich, Switzerland	S. Schmid
Centre of Poetics and Phonetics, University of Geneva, Switzerland	S. Vater

INTERNATIONAL SOCIETY OF PHONETIC SCIENCES: ADDRESSES

www.isphs.org

President:

Prof. Ruth Huntley Bahr, Ph.D.
President's Office:
University of South Florida
Dept. of Communication Science and
Disorders
4202 E. Fowler Ave., PCD 1017
Tampa, FL 33620-8200
USA
Tel.: ++1-813-974-3182
Fax: ++1-813-974-0822
e-mail: rbahr@cas.usf.edu

Secretary General:

Prof. Dr. Mária Gósy, DSc
Secretary General's Office:
Kempelen Farkas Speech Research
Laboratory
Hungarian Academy of Sciences
Benczúr u. 33
H-1068 Budapest
Hungary
Tel.: ++36 (1) 321-4830 ext. 172
Fax: ++36 (1) 322-9297
e-mail: gosy@nytud.hu

Guest Editor:

Tomáš Duběda, Ph.D.
Institute of Translation Studies
Faculty of Arts
Charles University in Prague
Hybernská 3
110 00 Praha 1
Czech Republic
Tel.: ++420 221 619 531
e-mail: dubeda@ff.cuni.cz

Review Editor:

Prof. Judith Rosenhouse, Ph.D.
Swantech
89 Hagalil St
Haifa 32684
Israel
Tel.: ++972-4-8235546
Fax: ++972-4-8327399
e-mail: swantech@013.net.il

**the
Phonetician**

A Publication of ISPhS/International Society of Phonetic Sciences

ISSN 0741-6164

Number 96 / 2007-II

Contents

From the President	6
From the Editor	6
Articles and Research Notes	
On the objectivity of prosodic phrases <i>by Jan Romportl</i>	7
Birthday	
Martin Kloster-Jensen: 90 th Birthday <i>by Joachim Neppert</i>	20
Obituary	
Přemysl Janota (1926–2008) <i>by Marie Dohalská</i>	21
Phonetics Institutes Present Themselves	
Department of Cybernetics, University of West Bohemia, Pilsen, Czech Republic <i>by Jan Romportl and Josef Psutka</i>	23
Conference Reports	
The Sixth Foreign Language Phonetics Teaching Conference <i>by Włodzimierz Sobkowiak</i>	28
The 14 th Manchester Phonology Meeting <i>by Marleen Spaargaren</i>	30
XI th International Conference SPECOM'2006 <i>by Andrey Ronzhin</i>	35
Beszédkutatás 2007 (Speech Research 2007 National Conference) <i>by Tekla Etelka Gráczy & Viktória Horváth</i>	39
The 2 nd Czecho-Slovak Conference of the International Society of Phonetic Sciences <i>by Tomáš Duběda</i>	41
Book Reviews	
Bisol, L. & Brescancini, C. R. (eds.): Contemporary Phonology in Brazil <i>reviewed by Ben Hermans</i>	44
McClelland, C. W.: A Phonology of Tarifit Berber <i>reviewed by Judith Rosenhouse</i>	52

Revue Parole, Issues 41–42 and 43–44 <i>reviewed by Judith Rosenhouse</i>	54
Peters, J.: Intonation deutscher Regionalsprachen <i>reviewed by Verena Teschke</i>	58
Sara, S. I. & Mauck, S. M. (eds. and trans.): Yehuda ben David Hayyuwj, Tanqiyt': A Treatise on Hebrew Vowels, Translation, Notes and Analysis <i>reviewed by Tamar Zewi</i>	61
Gardner Bonneau, D. & Blanchard, H. E. (eds.) Human Factors and Voice Interactive Systems <i>reviewed by Gábor Olaszy</i>	64
Meetings, Conferences and Workshops	68
Call for Papers	87
Instructions for Book Reviewers	87
ISPhS Membership Application Form	88
News on Dues	89

FROM THE PRESIDENT



I hope that you are enjoying the new format of the *Phonetician*! It takes a little time to get used to looking for it online, but it is nice to have the possibility of including color graphics and pictures in our newsletter. What I like most about the *Phonetician* now is that we are using a variety of editors to publish it. This method really highlights the international flavor of our publication. Having different editors allows us to tap into the talents of more members of our Society and gives all of us greater insight into what is happening in phonetics world-wide.

This issue provides a unique picture of phonetics in eastern Europe. The article, conference reports and other news are representative of that region. It is exciting to read about what is going on in these universities and institutes. Dr. Tomáš Duběda has done an excellent job of representing the work that is going on here. He is to be congratulated on putting together an exciting issue. Appreciation is also expressed to all of his contributors and I hope that we will continue to receive reports from this region.

Once again, I want to encourage YOU to get involved. Our readers want to know what is going on in your laboratory or institute. Not only does this advertise the work of colleagues in your corner of the world, but it can also encourage collaboration among scientists working in similar areas, who did not previously know about each other's work. After all, one of the missions of our Society is to educate and encourage researchers and to increase collaborations among scientists. I look forward to our new use of different editors as a way to increase our visibility and to encourage researchers around the world. Won't you please be a part of this exciting new venture?

FROM THE EDITOR



As the guest editor of this volume, I am glad to offer the members of ISPhS – and other readers as well – their regular collection of scientific information and news in the domain of phonetic sciences. To be honest, the word “regular” has shifted away from its canonical meaning with the last editions of *The Phonetician*, but we are now making up for this delay at a rapid pace.

After Trier, Budapest and Prague, the “itinerant editorship”, which is the new policy of *The Phonetician*, will be an occasion for the bulletin to come to us from other interesting places of the world.

I wish you much joy in reading *The Phonetician*, and express warm thanks to all those who have contributed to this number.

ARTICLES AND RESEARCH NOTES

ON THE OBJECTIVITY OF PROSODIC PHRASES

Jan Romportl

Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia
Univerzitní 8, Pilsen, 306 14 Czech Republic
rompi@kky.zcu.cz, Tel: +420 377 632 533, Fax: +420 377 632 502

Abstract

Objective annotation of prosodic phrases in a corpus for a text-to-speech system is an important issue due to its influence on the naturalness of synthesised speech. The paper discusses drawbacks of common ways of prosodic phrase annotation and proposes a concept of prosodic phrases defined by a maximum likelihood estimation over results of many parallel subjective annotations. Validity of this method is analysed in terms of agreement among the subjects using Cohen's and Fleiss' kappa measures and heuristically modified relative agreement.

Keywords

Prosodic phrase, objective annotation, speech corpus

1.0 Introduction

A text-to-prosody (TTP) system as a subsystem of a text-to-speech (TTS) system can be conceived and developed in terms of a machine learning (ML) paradigm. Such a conception, however, requires the existence of suitable training and testing databases covering desired prosodic phenomena. In this case, what does "suitable" mean? How much of such data do we need? And, most importantly, *who* can prepare such data?

I will try to explain my view on these questions, which can be easily classified as rather pragmatic – that is to say, I will hold the view that those data can be successfully prepared with very modest a priori phonetic knowledge and I will demonstrate this on the case of prosodic phrase and semantic accent annotation in a corpus for the Czech unit-selection TTS system, ARTIC. It is, nevertheless, very important to note that I do not deny the importance of phonetic knowledge per se – I only assert that in this particular situation, such knowledge is not essential.

The concept of *prosodic phrase* basically corresponds to a traditional view or to what is meant by the term "phonemic clause" (or "discourse segment") in Czech literature (Palková, 1974), i.e. a phonetic unit which underlies the perception of a certain level of rhythmical qualities in language. A prosodic phrase is mainly delimited by the acoustical

features of its boundaries and it can also contain an “intonation peak”. However, as Palková discusses (ibid.), there is no empirical evidence supporting any stronger assumption about the presence/absence of an intonation peak or their number in a Czech utterance.

We further assume that a speaker may emphasise any number of words by acoustic means to express (perhaps even unintentionally) their prominence in comparison with other words. The acoustic prominence of a word can deliver various kinds of information: it can either help structure an utterance and delimit phrase boundaries, or it can modify the semantics and pragmatics of an utterance. The first type of acoustic prominence is automatically realised at the end of a phrase. The second type, called *semantic accent*, can be realised anywhere within an utterance and its usage is semantically functional – it often plays an important role in the articulation of topic-focus.

Prosodic phrase and semantic accent usage can be illustrated by several examples. Prosodic phrase boundaries are designated by “/”, words in italics are with the semantic accent:

- “tito živočichové / jsou velice inteligentní / ale také pomalí” – “these animals / are very intelligent / but also slow”
- “i hráčům *jiných* sportů / jdou šipky dobře” – “players of *other* sports / are also good in darts”
- “podle *Iráku* / jich bylo *pouze* osm” – “according to *Iraq* / there were *only* eight of them”

2.0 Aspects of prosodic phrase annotation

The prosodic phrase annotation process I will discuss here has one specific goal: to allow the designed TTP system based on ML techniques to produce prosodically natural speech in terms of phrasing. Hence, this process is not primarily focused on investigating the nature of prosodic phrases but the results can help identify them. The aspects of semantic accent annotation are analogous to the following aspects of prosodic phrase annotation.

2.1 Speech synthesis requirements

The first aspect of the prosodic phrase annotation is determined by various demands posed by speech synthesis techniques – namely by the unit-selection concatenative synthesis algorithm.

This algorithm is based on selection of speech units according to their classification into various relevant, mostly structural categories. A synthesized utterance is represented as a target sequence of units and their categories, and the synthesis algorithm tries to find units from the speech segment database matching the target sequence as closely as possible.

This means that every unit in the database must be described (annotated) by values of all the categories used in the algorithm.

Since one of the categories is based on prosodic phrasing (e.g. position of a unit in a prosodic phrase, etc.), prosodic phrases must be determined for every sentence of the source speech corpus whose units appear in the speech segment database. Most current unit-selection TTS systems try to make use of as much data as possible (so as to achieve the most natural sounding speech), thus speech segment databases usually consist of thousands of sentences. In the case of our TTS system, the database consists of 10 000 Czech declarative sentences (taken from newspaper texts) recorded in a studio by a male professional speaker. All of the sentences were annotated with prosodic phrase boundaries.

It should be mentioned that the difficulty of prosodic phrase placement strongly depends on the type of the text in the corpus and on the speaker. If the text is uttered very rhythmically and affectively (such as heard in a good actor's performance), it is easier to find prosodic phrases than in a neutral and intonationally "flat" speech. Due to the constraints posed by the TTS techniques our speech corpus is of the latter kind. Moreover, prosodic phrase boundaries in Czech speech often tend to be more vague and ambiguous than in English.

2.2 Automatic annotation

Due to the large amount of data to annotate (for a single voice – and most TTS systems offer many voices) it seems to be inevitable for machines to replace human annotators. Another, and perhaps more important reason is consistency in transcription. It is extremely difficult (or even impossible) for a human annotator to maintain consistent perception and annotation of such phenomena throughout corpora including thousands or tens of thousands of sentences.

The idea is to manually (i.e., by humans) designate prosodic phrase boundaries and semantic accents in a reasonable sub-part of the whole corpus (250 sentences in our case) and then use ML pattern recognition techniques to automatically and consistently extend this annotation to the whole corpus (10 000 sentences in our case).

This method, however, imposes even stronger demands on consistency of the manual annotations. Should there be discrepancies between acoustic/textual cues and phrase boundary judgements in training and testing data, reliability of the ML classifier may seriously decrease.

2.3 Prosodic phrases as theoretical entities

Prosodic phrases are what we define them to be. Their ontological status is the same as that of other theory-based entities. Only through theories do we know what a prosodic phrase is.

“Theories are nets cast to catch what we call ‘the world’: to rationalize, to explain, and to master it. We endeavour to make the mesh ever finer and finer.” (Popper, 2002, p. 38)

Some theories define prosodic phrases by their boundaries, realised as particular f_0 or duration movements. These definitions are undoubtedly useful for certain purposes, but due to their reductionist form they lack one important attribute – *function*. I believe that defining prosodic phrases (and not only them) through their function is much more epistemologically valuable, and that the goal of a TTP system is to generate prosody with a proper function (from the point of view of a listener), no matter its actual form.

We can say that the function of a prosodic phrase is: 1) to create one layer of the rhythmical structure of an utterance (this rhythmical layer is hierarchically and structurally higher than the level of prosodic words and lower than the level of utterances); 2) to help a listener reconstruct the underlying nonlinear structure of the utterance. This means that a prosodic phrase corresponds to a continuous segment of an utterance where a single instance of the function (1) and a disposition for the function (2) are prosodically realised (i.e. realised by means of a prosodic form). On one hand this definition brings “epistemological sense” into looking for prosodic phrases, on the other hand it introduces subjectivity and hence inconsistencies.

This brings us to the following problem: the TTP system should generate functionally proper prosody, thus the whole corpus must annotate prosodic phrases in the aforementioned functional-perceptual sense. The training/testing datasets also must be manually annotated in the same sense. Such manual annotation is likely to be inconsistent – and this is indeed unwanted.

What is and what is not a prosodic phrase (or a prosodic phrase boundary) from a scientific point of view can be formulated and decided by an empirical theory. However, it seems that this theory can only rely on either subjective judgements about perception, rhythm, syntactical disambiguation, etc., or reductively on non-functional cues, such as f_0 movements, etc. The former case is scientifically problematic because empirical facts are substituted by subjective beliefs, the latter lacks functional relevancy.

If a trained phonetician manually annotates a part of a speech corpus, he might be consistent in his subjective judgements (because of his training) and might strictly obey the principles of a particular theory about prosodic phrases, but one can immediately find a group of people (non-specialists) who will disagree with a significant number of phrase boundaries placed by this expert in phonetics. This means that such a prosodic phrase annotation does not represent prosodic phrases where they really are, but where one person thinks they are. This situation actually looks like the particular phonetic theory “knew” where the prosodic phrases really are and the phonetician either “hits” or “misses” them. We could also consider three or five or more phoneticians doing this job as a team – they would try to make best of their experience and knowledge of the theory, they would discuss where the theory posits phrase boundaries and they would eventually settle on some mutually agreeable decision. Still, there would be no way to see how close to the “real” phrase placement their decision is. So, we must ask ourselves a question: what is the nature of the empirical statements (about prosodic phrases) of such a theory?

Of course if we want to discover the nature of prosodic phrases from within the language system, it is perfectly correct to posit theoretical features of prosodic phrases a priori and then test them on real speech data. The only problem in this particular case is that the testability of these units is rather questionable, as I have already argued.

However, if our primary goal is to know objectively and exactly where the prosodic phrases in a particular speech corpus are, we can describe the prosodic reality by a different, more pragmatic theory. After all, such objective knowledge can be of great usefulness for testing other prosodic theories.

The theory I am proposing here to define what we call a “prosodic phrase” (at least in the Czech language and at least for the sake of what I have discussed above) comprises following assumptions:

1. Every normal speaker/listener (native speaker/listener) has an intuitive sense of rhythm in speech. The purpose of this rhythm is to help perceive and structure utterances.
2. We can suppose that speech rhythm is constituted by specific units, which are, on a certain structural level, called prosodic phrases. This is an important piece of knowledge we are borrowing from other theories.
3. There is a probabilistic causal relationship between the presence of the boundary of a prosodic phrase and the intuitively (subjectively) conditioned conscious designation of this boundary by a listener.
4. Empirical facts are statements about behaviour of listeners – a listener either asserts that there is a prosodic phrase boundary at a certain place in speech or asserts that there is not.
5. If there is a statistically relevant number of empirical facts from independent listeners describing the same portion of speech, a model of an *objective annotator* can be created. The objective annotator is a maximum likelihood estimation over the empirical facts.
6. A prosodic phrase is what is designated by the objective annotator.

The nature of prosodic phrases based on these assumptions is entirely clear, testable and reproducible. It is quite likely that there would be differences between the “opinion” of the objective annotator and the opinion (perhaps collective) of the aforementioned phonetic experts. Although it may be interesting to analyse such differences, one must keep in mind that, metaphorically speaking, it is comparing two different (theoretical) worlds without clear bridging links or principles.

It might seem rather vague to use the term “statistically relevant number”, but we can define this number more precisely as the number of listeners which satisfies the condition that an objective annotation created over this set of listeners equals to the objective annotation created over this set extended by one more arbitrary listener.

We can go even further in exploiting the aspects of our task described in the sections 2.1 and 2.2 and take into account the automatic annotation by these assumptions:

7. The *objective machine annotator* is such a classifier set up by ML techniques which achieves the highest possible classification performance on the testing data prepared by the objective annotator.
8. Prosodic phrase is what is designated by the objective machine annotator.

This allows us to acquire stable and consistent prosodic phrase annotation of speech corpora of an arbitrary size without doubts (towards or in the sense of the theory) about its objectiveness. There is one more question: can we do all these things also for semantic accents?

3.0 Experiments

The annotation process with the aspects described in Section 2.0 has been based on two large-scale listening tests – I will further denote them as Test 1 and Test 2.

3.1 Listening tests

The listening tests were organised on the client-server basis using a specially developed web application. We used the speech corpus which the text-to-speech system ARTIC (Matoušek – Romportl, 2007) is based on. The corpus was very carefully recorded in a studio by an experienced male speaker (the choice of the speaker was made in consultation with two experts from the Institute of Phonetics, Charles University in Prague) who had been instructed to read isolated sentences naturally, yet avoiding any expressiveness. The speaker did not know that the recorded sentences also would be used for the phrasing analysis. The way the corpus has been recorded (i.e., the type of recorded speech) obviously influences the scope of linguistically relevant findings of the research – therefore the relevance of the quantitative results presented here is limited to the aforementioned speech domain; however, the methods we have used are definitely not limited to this data.

3.1.1 Test 1

In Test 1, we randomly selected 100 sentences from the corpus and loaded them together with their orthographic transcriptions into the web application. Potential test participants were selected among university students from all faculties (with a special focus on students of linguistics). When they finished the listening tests, they were financially rewarded (so as to increase their motivation). The participants could do all of the work from their homes without any personal contact with the test organisers – we have thus undertaken various measures to detect possible cheating, carelessness or misunderstandings.

The participants were instructed to listen to the sentence recordings very carefully and subsequently designate words where they are sure there is a phrase boundary and words where they feel there might be a phrase boundary (i.e., these two cases were carefully distinguished). Prior to the test itself the participants had been briefly familiarised with the

background of the problem and in this tutorial they listened to several training samples which showed possible phrasing demonstrations. It is, however, very important to note that we intentionally did not want to make almost any a priori assumptions about phrase boundary qualities or behaviour. We wanted to create a “notion of prosodic phrase” in the participants and let them designate whatever subjectively fulfilled this notion (cf. Section 2.3).

We eventually received correctly finished tests from 103 participants (the total number of students who took part in these tests was 174, some of the students had not finished their tests, some of them had not even started, and there were also several apparent cheating attempts), which provided a robust observation set for further evaluation. Several quantitative facts about the Test 1 are in Table 1.

Table 1. Quantitative facts about Test 1 and Test 2

	Test 1	Test 2 (Part 2)
Finished tests	103	99
Participants with phonetic education	25	19
Average time spent on one test	92 min	168 min
Average number of sentence replays	2.33	2.25
Average number of sessions per user	3.10	5.14
Total number of sentences	100	150
Total number of word tokens	1063	1531
Total length of speech	≈ 508 s	≈ 741 s

3.1.2 Test 2

Test 2 (which was carried out 3 months after Test 1) consisted of two parts (hereafter denoted as Part 1 and Part 2). Part 1 focused on finding the semantic accents in sentences where the prosodic phrase boundaries were already given. The same sentences from Test 1 were used again and the participants had been instructed to listen to these sentences very carefully and subsequently designate words where they perceived semantic accent. The textual form of the sentences was displayed together with the a priori prosodic phrases acquired from the objective annotation based on Test 1. The participants had to accept this phrasing and adapt their semantic accent assignment accordingly. Part 1 also served as a “tutorial” for Part 2 since the participants could infer how to annotate the prosodic phrases in Test 1.

Part 2 was actually a combination of Part 1 and Test 1: we selected another 150 sentences from our corpus and the participants were again instructed to listen to the sentence recordings and designate the semantic accents. However, in this part, the task was also to designate words with perceived prosodic phrase boundaries (cf. Section 3.1.1).

The quantitative facts about Test 2 can be compared with Test 1 again in Table 1.

3.2 Objective annotation

We can now describe the problem of modelling annotation based on many independent observations on a more abstract and formal level:

Let X be a random process defined as $X = \{X_t : t \in T\}$, where $T = \{1, 2, \dots, n\}$ is a set of time points respective to the ordinal numbering of words in the test sentences (i.e., the first word in the first sentence has $t = 1$, the second word in the first sentence has $t = 2$, and so on), and X_t are random variables which hold $X_t = 1$ if and only if the t -th word finishes a prosodic phrase, and $X_t = 0$ otherwise. Exactly the same can be done for the semantic accents, such a random process is analogous to X and will be denoted as Y . We assume that the random processes X and Y are mutually independent.

Now let the test participants be numbered by the set $J = \{1, 2, \dots, m\}$, i.e., the first participant has $j = 1$, the last one has $j = m$. We can define m random processes $O^{(1)}, \dots, O^{(m)}$ representing the participants' responses (observations, empirical facts) such that $O^{(j)} = \{O^{(j)}_t : t \in T\}$, where t has the same meaning as for the process X , and $O^{(j)}_t$ are random variables which hold $O^{(j)}_t = 1$ if and only if the j -th participant asserts that the t -th word finishes a prosodic phrase, and $O^{(j)}_t = 0$ if and only if the j -th participant does not assert that the t -th word finishes a prosodic phrase.

Our goal can now be re-formulated as follows: knowing the observations $O^{(1)}, \dots, O^{(m)}$, we want to estimate the hidden trajectory of the process X which best satisfies the given observations. This can be determined analogically for the process Y . For the sake of clarity, I will speak further in the text only about the process X , assuming that everything which holds for it, also holds for the process Y . It is supported by the fact that the two variants of the answers on the phrase boundary presence/absence (i.e., “boundary certain” and “boundary maybe”) were treated equally – this was based on the assumption that if the “statistically relevant” number of participants think that there *might be* a phrase boundary at the given place, it *really is* there. The reason for allowing two levels of certainty from the participants' view was mainly due to the experience that if a listener is really not sure, he answers randomly – and this can be avoided by the “maybe” variant. The difference between these two variants is utilised in the participants' agreement calculation (see Section 3.3.2).

The aforementioned goal of the hidden trajectory estimation can be transformed into the problem of finding the most likely model parameters given the observed data – a *maximum likelihood* approach (cf. Section 2.3). I will not describe this method here because it involves some mathematics, and I have described it elsewhere (Romportl, 2008). In any case, the result of this method is the objective annotation of 250 sentences with both prosodic phrases and semantic accents

3.3 Validity of the objective annotation

The validity of the objective annotation can be interpreted as a quantitative measure of inter-participant agreement. This way we can test the validity of the assumptions listed in Section 2.3. If the measure of inter-participant agreement is too low, it will suggest that one or more assumptions should be modified.

3.3.1 Kappa measures

The agreement between two test participants can be measured by means of a statistical correlation. However, if one of the answers in the test is significantly more frequent than the other, two participants can often agree just by chance and the correlation is thus relatively high, therefore misleading. For example, if the first participant designates phrase boundaries where they really are and the second participant asserts there is no phrase boundary in the corpus, their correlation still will be relatively high because they will often “agree” on the non-boundary words which are more frequent than the boundary ones.

Such influence of the agreement by chance can be eliminated by using Cohen’s and Fleiss’ kappa measures (κ_C , κ_F). Cohen’s kappa (Cohen, 1960) is a scalar value measuring agreement between two test participants; Fleiss’ kappa (Fleiss, 1971) expresses agreement among more participants at once.

We calculated κ_F for Test 1 and Test 2 separately and then for the whole set of 250 sentences in two variants – including and excluding words followed by a pause (a phrase boundary with a pause is much easier to detect). This was calculated for semantic accents too. As κ_C measures only mutual agreement between two participants, we calculated it for every pair of the participants and then presented it as the average value. The results are displayed in Table 2. Moreover, κ_C can be also calculated for every participant paired with the objective annotator – this is summarised in Table 3.

Table 2. Values of Fleiss’ and Cohen’s kappa. $E\{\kappa_C\}$ is the average value for all pairs of the participants, $D\{\kappa_C\}$ is the variance.

	Prosodic phrases				Semantic accents		
	Whole set	Whole set excl. pauses	Test 1	Test 2	Whole set	Test 1	Test 2
κ_F	0.5790	0.4171	0.4542	0.6636	0.1283	0.1325	0.1201
$E\{\kappa_C\}$	0.5837	0.4293	0.4632	0.6710	0.1271	0.1417	0.1259
$D\{\kappa_C\}$	0.0068	0.0154	0.0180	0.0083	0.0052	0.0048	0.0069
max κ_C	0.7669	0.8179	0.8538	0.8929	0.7690	0.7910	0.7875
min κ_C	0.1718	0.0927	0.0801	0.1978	0.0081	0.0072	0.0010

Table 3. Values of Cohen’s kappa for the participants paired with the objective annotator. $E\{\kappa_C\}$ is the average value, $D\{\kappa_C\}$ is the variance.

	Prosodic phrases				Semantic accents		
	Whole set	Whole set excl. pauses	Test 1	Test 2	Whole set	Test 1	Test 2
$E\{\kappa_C\}$	0.7100	0.5637	0.6182	0.7729	0.2596	0.2977	0.2574
$D\{\kappa_C\}$	0.0061	0.0086	0.0192	0.0058	0.0069	0.0061	0.0078
max κ_C	0.8242	0.8419	0.8173	0.9013	0.4572	0.4610	0.4638
min κ_C	0.3854	0.1301	0.0000	0.3780	0.0894	0.0613	0.0000

3.3.2 Heuristically modified relative agreement

In spite of the kappa measures (both Cohen’s and Fleiss’) being the chance-corrected measure of agreement, their usage and more importantly their interpretation is often rather problematic (Maclure – Willett, 1987). It is thus advisable to supplement them with another, more informed quantitative criteria.

Although it might seem that mere relative agreement between two participants is not a very good choice, I am convinced that if this simple measure is slightly heuristically modified, it can provide a statistical tool which takes into account information about behaviour of the participants and relevancy of various types of answers.

I therefore propose two types of relative agreements for prosodic phrase boundaries:

- The agreement of a pair of participants is calculated as the number of cases in which both participants chose the same answer, divided by the total number of answered cases. More formally: the agreement $A_1(i, j)$ between the participants i and j is defined as

$$A_1(i, j) = \frac{\sum_{t \in T} f_{ij}(t)}{n},$$

and

$$f_{ij}(t) = \begin{cases} 1 & \leftrightarrow (\rho(o_t^{(i)}) = \rho(o_t^{(j)})) \\ 0 & \leftrightarrow (\rho(o_t^{(i)}) \neq \rho(o_t^{(j)})) \end{cases},$$

where $\rho(x)$ is integer rounding of x (will be explained later). The overall average agreement of this type is then given as

$$A_1 = \frac{\sum_{i, j \in I, j > i} A_1(i, j)}{\frac{1}{2}m^2 - m}.$$

- The agreement of a pair of participants is calculated as the number of all cases in which *both* participants chose the *positive* answer, divided by the number of cases in which *at least one* of these two participants chose the *positive* answer. In this way, the agreement calculation is motivated by the heuristic knowledge that a vast majority of cases agreeing by chance involve negative answers; moreover, the

agreement on *absence* of a prosodic phrase has epistemologically “lower” modality than the agreement on its *presence*. Again, formally we can write

$$A_2(i, j) = \frac{\sum_{t \in T} (f_{ij}(t) \cdot c_{ij}(t))}{\sum_{t \in T} c_{ij}(t)}$$

where

$$c_{ij}(t) = \begin{cases} 1 \leftrightarrow (O_t^{(j)} = 1 \vee O_t^{(i)} = 1) \\ 0 \leftrightarrow (O_t^{(j)} = 0 \wedge O_t^{(i)} = 0) \end{cases}$$

I have mentioned in Section 3.2 that the participants actually had three choices when answering whether a particular word bears a phrase boundary: “yes”, “no” and “maybe”. Although the “maybe” variant is treated as “yes” in the process of the objective annotation estimation, the way how it is interpreted in the process of evaluation can influence the result of the evaluation. We can introduce three methods of interpretation of the “maybe” variant:

- **M1**: There is no difference between the “maybe” and the “yes” variant. This means that $O_t^{(j)} = 1$ in both cases: the j -th participant designates the t -th word as “certainly with phrase boundary” or he designates it as “maybe with phrase boundary”.
- **M2**: The “maybe” variant is ignored, i.e. $O_t^{(j)} = 0$ anytime the j -th participant designates the t -th word as “maybe with phrase boundary”.
- **M3**: The “maybe” variant has a different value than the “yes” variant: $O_t^{(j)} = 1$ in case the j -th participant designates the t -th word as “certainly with phrase boundary” and $O_t^{(j)} = 0.6$ in case the j -th participant designates the t -th word as “maybe with phrase boundary”. It is the “maybe” variant here for which I have defined the operator $\rho(x)$ for integer rounding. For this method it is also necessary to slightly modify the following equation:

$$c_{ij}(t) = \begin{cases} 1 \leftrightarrow (O_t^{(j)} = 1 \vee O_t^{(i)} = 1) \\ 1 \leftrightarrow (O_t^{(j)} = 0.6 \wedge O_t^{(i)} = 0.6) \\ 0 \leftrightarrow \text{otherwise} \end{cases}$$

The combination of these three methods with the functions A_1 and A_2 gives us six ways how to compare a pair of participants. Table 4 summarises values of these heuristically modified relative agreements.

Table 4. Heuristically modified relative agreement between the test participants. The most relevant and informative values are bold.

	M1		M2		M3	
	A_1	A_2	A_1	A_2	A_1	A_2
Test 1	0.81	0.41	0.86	0.41	0.81	0.56
Test 2	0.93	0.43	0.95	0.43	0.93	0.74
Whole set excl. pauses	0.73	0.37	0.76	0.37	0.73	0.53

4.0 Discussion and conclusion

On the basis of the values presented in Table 2 and 3 it is clear that the validity of prosodic phrases based on the assumptions from Section 2.3 is well supported because the inter-participant agreement is relatively high. The Fleiss' kappa value is very similar to the values for English presented in recent studies (Mo et al., 2008).

The average agreement between the participants and the objective annotator is also high and it suggests that prosodic phrases defined via the objective annotator are not a mere formal construct, but they maintain a very strong link with human perception. This is a very important conclusion.

Semantic accents, on the other hand, are significantly more difficult to test. Thus, their existence in terms of their definition from Section 1.0 is very questionable. Therefore, the estimation of the process Y should probably not be called the objective annotation (of semantic accents). It is the most stable and objective annotation on which we can base the responses of the test participants (i.e., on the empirical facts about their assessment of semantic accents), but the actual responses acquired in Test 2 are apparently rather chaotic and too inconsistent. Although we can subjectively agree that there really is “something” in many utterances that “sounds emphasised”, we will need a different theory to be able to capture such phenomena objectively.

Although the kappa measures are a good quantitative indicator, the heuristically modified relative agreement is easier to interpret when creating overall judgement about the acquired data: high agreement calculated by the function A_1 with the method $M1$ significantly decreases when using A_2 , which implicates that the participants evidence strong agreement on the absence of phrase boundaries. The most informative value about the agreement on the presence of boundaries, taking the “maybe” variant in consideration is given by the combination of A_2 and $M3$.

I would also point out the differences between Test 1 and Test 2. The different values of the agreement measures are most likely caused by two reasons: 1) in Part 2 of Test 2, the participants had already passed the annotation of semantic accents from Part 1, so they had already been familiarised with the phrase objective annotation from Test 1. This means they could acquire better implicit understanding of the phenomenon of prosodic phrases; and 2) it probably makes a difference whether the participants designate phrases and semantic accents separately or at the same time. This will be the focus of future investigations designed to evaluate the extent to which these two reasons could have influenced the results.

The current state of development and performance of the objective machine annotator is presented elsewhere (Romportl, 2010). The classifier is based on artificial neural networks and is able to designate prosodic phrases in the whole speech corpus significantly better than an average human annotator.

Considering these results, I think the proposed pragmatic theory of prosodic phrases can be characterised as acceptable. There are still many questions to be answered, such as the appropriateness of the criterion for the “statistical relevance” of the number of empirical

facts, or possible differences in the objective annotation given different sets of test participants. Answering these questions, however, will demand data based on a new, specially and carefully designed listening test.

Acknowledgements

Support for this work was provided by the Ministry of Education of the Czech Republic, project LC536.

References

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76.
- Maclure, M. & Willett, W. C. (1987). Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology*, 126, 161–169.
- Matoušek, J. & Romportl, J. (2007). Recording and annotation of speech corpus for Czech unit selection speech synthesis. In *Proceedings of TSD 2007*, Lecture Notes in Artificial Intelligence, vol. 4629 (pp. 326–333). Berlin–Heidelberg: Springer.
- Mo, Y. & Cole, J. & Lee, E.-K. (2008). Naive listeners' prominence and boundary perception. In *Proceedings of Speech Prosody 2008* (pp. 735–738). Campinas, Brazil.
- Palková, Z. (1974). *Rytmičká výstavba prozaického textu* (with English resume: The rhythmical potential of prose). Prague: Academia.
- Popper, K. R. (2002). *The Logic of Scientific Discovery*. London: Routledge.
- Romportl, J. (2008). *Zvyšování přirozenosti strojově vytvářené řeči v oblasti supra-segmentálních zvukových jevů* (Improving Naturalness of Machine-Generated Speech on the Suprasegmental Level). Ph.D. dissertation, Department of Cybernetics, University of West Bohemia, Pilsen.
- Romportl, J. (2010). Automatic prosodic phrase annotation in corpus for speech synthesis. In *Proceedings of Speech Prosody 2010* (to be published in May 2010). Chicago IL, USA.

BIRTHDAY

Martin Kloster-Jensen: 90th Birthday



Prof. Dr. Martin Kloster-Jensen, born on 23rd of February 1917 in Norway, has recently celebrated his 90th birthday. From 1961 to 1991 he served our society in various capacities, for example as Secretary General and as President. He is one of ISPhS' Honorary Vice Presidents, was awarded the status of Fellow in 1979, and he has been a Laureate since 1983. Further details on his life, his work, his awards, and his publications are available in *The Phonetician* Number 86 (2002-II).

We congratulate him on the occasion of this outstanding birthday and express our best wishes for many happy returns.

Joachim Neppert
Am Sooren 51
D-22149 Hamburg

OBITUARY

Přemysl Janota (1926–2008)



Přemysl Janota, professor emeritus of the Prague Institute of Phonetics, died on September 25, 2008, leaving behind him a life's work in phonetic sciences.

The interest of P. Janota (born on May 10, 1926) in phonetic and linguistic problems, as well as acoustics, started during the War, while he attended secondary school. His interest in the Dutch language dates from the same period. After the War, he studied phonetics, general linguistics, English and philosophy at the Faculty of Arts, Charles University. He obtained the doctor's degree in phonetics and linguistics in 1950, with a thesis entitled *On measuring speech dynamics*.

Přemysl Janota first worked as a phonetician and linguist in the Prague Institute of Logopaedics. At the same time, he taught Dutch at the Faculty of Arts. In the field of audiology, he developed an original therapeutic method, and was the first in the former Czechoslovakia to work up, both linguistically and technically, a child-oriented method of audiometry.

In 1957, Přemysl Janota returned to the Institute of Phonetics, where he taught courses devoted to theoretical and practical questions of phonetic methodology, speech acoustics, general phonetics, phonetics of Czech, German, Dutch, and later English. While in charge of the Institute's laboratory, he designed and constructed more than 30 devices, many of which were appreciated abroad. The most cited instrument is undoubtedly "Janota's segmentator". Another important device designed by P. Janota was the first vowel synthesizer in Czechoslovakia.

In the 1960s, P. Janota investigated individual properties of the speech signal and the perceptual characteristics of vocalic sounds. The most important outcome of this activity

was the book *Personal characteristics of speech* (1967). At the same time, he devoted himself to intensive research on the speech of hearing-impaired children and adults.

The versatile personality of P. Janota included a remarkable pedagogical talent. Several generations of students have experienced his lectures and seminars.

During his professional career, P. Janota participated in a number of research projects and collaborated with different institutions (e.g., Institute of Communication Engineering, Phonetic Laboratory of the Academy of Sciences, Technical University Dresden). He gave lectures in the Netherlands, England, Sweden and Norway, to mention just a few countries, and attended a large number of international congresses and meetings. Thanks to his universal erudition and language talent, he was asked more than once to chair round-table discussions. He was one of the vice-presidents of the ISPhS.

Přemysl Janota was appointed professor emeritus of the Charles University in 1995.

Marie Dohalská
Charles University in Prague, Czech Republic
marie@dohalska.cz

PHONETICS INSTITUTES PRESENT THEMSELVES

Department of Cybernetics, University of West Bohemia Pilsen, Czech Republic

The Department of Cybernetics has been a member of the science and research community in Pilsen (Plzeň) since 1960; first as a part of the Faculty of Electrical Engineering at the Institute of Technology in Pilsen. The department received its present modern face in 1990 when it became one of the five departments of the newly founded Faculty of Applied Sciences. Today, the Faculty of Applied Sciences is the most dynamic part of the University of West Bohemia, having also the highest research potential (UWB currently hosts 11 faculties, schools and centres with almost 20 thousand university students and staff members).

There are three sections (or laboratories) at the Department of Cybernetics: Automatic Control, Information and Control Systems, and Artificial Intelligence. It is the Artificial Intelligence Section (AIS) with approximately 40 researchers and Ph.D. students, where research in speech science and technology is intensively conducted.

Speech in relation with machines has been the main field of interest of AIS since 2000 (although speech research had been conducted there before 2000 as well). Today, AIS is the research body with the most complex activities in this area in the Czech Republic. In spite of being very interconnected and synergic, this research can be roughly divided into five streams:

- Automatic Speech Recognition (ASR)
- Text-to-Speech Synthesis (TTS)
- Speaker Identification and Verification (SIV)
- Spoken Dialogue Systems (SDS)
- Visual Data Processing (VDP)

Automatic Speech Recognition

The Department of Cybernetics represents state-of-the-art research of large-vocabulary continuous speech recognition (LVCSR) of Czech and other Central and Eastern-European Languages. Apart from the development of real-time ASR applications (various large-vocabulary dictation systems, which are literally for every-day use on a regular PC), the bulk of the ASR research focuses on improvements in decoding performance, confidence measures, transparent speaker adaptation, larger vocabularies (currently 300k+ word forms, and in the near future aiming at 1M), more robust probabilistic language models, and keyword spotting techniques. The need for large vocabularies (in comparison with English) is especially felt in languages with very rich flexion, such as Czech and other Slavic languages.

An important period in the scientific development of the department was its intensive participation in the project MALACH (Multilingual Access to Large Spoken Archives),

funded by the National Science Foundation of the U.S.A. Main partners in this project were: Shoah Visual History Foundation (currently, “USC Shoah Foundation Institute for Visual History and Education”; the original aim of the foundation was to record testimonies of survivors and other witnesses of the Holocaust), Johns Hopkins University, IBM, University of Maryland, and Charles University in Prague (cf. <http://malach.umiacs.umd.edu/>). Our department was responsible for the development of very robust ASR systems for Czech, Slovak, Polish, Russian and Hungarian. These systems were then used for transcription of mostly very emotive and moving audiovisual interviews with survivors of the Holocaust, in their respective languages.

Another important project is being carried out in cooperation with Czech Television (the public service broadcaster in the Czech Republic) under the name “Elimination of the Language Barriers Faced by the Handicapped Watchers of the Czech Television” (ELJABR) and its most important contribution in the field of ASR was automatic subtitling of live TV broadcasting (either directly from the original acoustic track or from a shadow speaker). The pilot system for subtitling of parliament discussions has been tested in public broadcasting since November 2008. We have also tested the system for subtitling live ice hockey matches and for shadow speakers (Ircing et al., 2009; Pražák et al., 2006).

We are also involved in a large international project COMPANIONS funded by the European Commission’s Sixth Framework Programme. The aim of this project is to go beyond the state-of-the-art in man-machine communication by developing a virtual conversational “companion”. The Department of Cybernetics works on the ASR input of the Czech prototype of such a device.

Text-to-Speech Synthesis

The area of TTS synthesis has been investigated at our department since 1997 (Matoušek, 2001). The department has pioneered Czech concatenative TTS synthesis (system ARTIC) based on very large speech corpora processed by statistical methods, such as automatic segmentation by Hidden Markov Models (until then the Czech TTS systems had been using small, manually prepared corpora). This is exactly the trend we followed later on, which has allowed the department to develop the first Czech naturally sounding unit selection TTS system based on a corpus counting 5 000, and later more than 10 000 professionally recorded sentences (Tihelka, 2005).

Current research focuses on prosodic aspects of naturally sounding synthetic speech (Romportl, 2010 [this volume]), emotional speech synthesis (Zovato & Romportl, 2008), unit selection algorithm improvements (Tihelka & Romportl, 2009), voice conversion (Hanzlíček & Matoušek, 2008), large corpora preparation, recording and processing (Matoušek et al., 2008), development of the first Czech HMM speech synthesiser, and other related topics.

Recent projects include the aforementioned ELJABR and COMPANIONS. One particular aspect of the former project is the study of an accompanying audio track of synthesised speech for TV broadcasting; the latter one implements emotional speech synthesis in the virtual “companion”.

Speaker Identification and Verification

SIV is another important area of speech technology research with strong potential for practical application. The main focus in this area is the development of expert systems for speaker verification based on a hybrid generative-discriminative approach combining Gaussian Mixture Models and Support Vector Machines (Zajíc et al., 2008). The Department of Cybernetics has also served as an officially appointed expert in legal proceedings where the identity of a speaker has to be determined.

Spoken Dialogue Systems

SDS research and development is a multidisciplinary issue where areas of expertise other than speech technology also play an important role: natural language processing and generation, dialogue management, knowledge management and manipulation, machine inference, etc. Apart from obvious involvement of ASR and TTS in state-of-the-art speech dialogue systems (such as the aforementioned COMPANIONS), there are speech-related problems specific to dialogue systems. One of them is barge-in detection and handling, which has been successfully implemented in our new state-of-the-art statistically driven telephony dialogue system providing real-time information about train arrivals, departures, time schedules, prices, etc. (Jurčiček et al., 2008). This system benefits from a very robust ASR and dialogue models, and can handle spontaneous phone calls from public users almost in the same way a human operator would do.

The Department of Cybernetics has been operating the University VoiceXML Telephony Information Systems for students and employees of the University of West Bohemia since 2000. This system, using ASR and TTS interfaces, provides information, for example, about internal telephone numbers or exam results, and allows students to register for exams.

Visual Data Processing

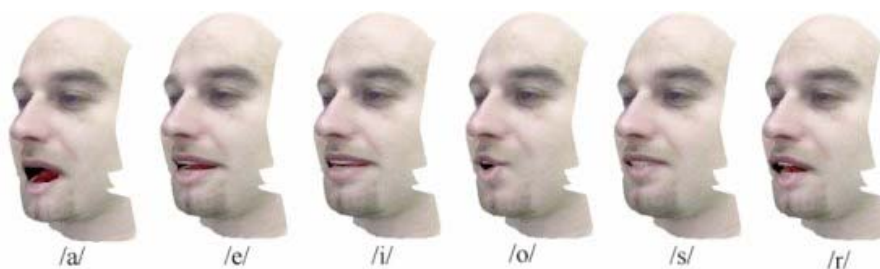
Although the VDP area usually involves various image processing tasks not directly connected with speech, there is a research team at the Department of Cybernetics which develops systems for audiovisual speech synthesis and recognition (and synthesis and recognition of sign language as well).

The idea of audiovisual speech synthesis and recognition is that visual information (usually about articulation movements) can significantly increase the successfulness of a communication process, mainly in noisy environments. A special use for these processes is pure visual communication for hearing impaired people by means of lip reading, which can be supported by sign language.

Audiovisual speech synthesis suitable for lip reading is implemented in our “Talking head” (see the pictures). This Talking head won the Golden Lips Award at Interspeech 2008 in Brisbane for the best articulation visualisation in English.



A “Talking head” for audiovisual speech synthesis.



Another “Talking head” demonstrating visual samples of articulation of several Czech phonemes.

References

- Hanzlíček, Z. & Matoušek, J. (2008). On using warping function for LSFs transformation in a voice conversion system. In *Proceedings of ICSP '08*, vol. 3. Beijing, China.
- Ircing, P., Psutka, J. V., Psutka, J. (2009). Using morphological information for robust language modeling in Czech ASR system. *IEEE Transactions on Audio, Speech and Language Processing*, 17, May issue, 840–847.
- Jurčíček, F., Švec, J., Müller, L. (2008). Extension of HVS semantic parser by allowing left-right branching. In *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. 4993–4997). Las Vegas, U.S.A.
- Matoušek, J. (2001). *Syntéza řeči z textu s využitím statistického přístupu k automatickému vytvoření databáze řečových jednotek*. Ph.D. dissertation, Department of Cybernetics, University of West Bohemia, Pilsen.
- Matoušek, J., Tihelka, D., Romportl, J. (2008). Building of a speech corpus optimised for unit selection TTS synthesis. In *Proceedings of LREC 2008*. Marrakech, Morocco.

- Pražák, A., Psutka, J. V., Hoidekr, J., Kanis, J., Müller, L., Psutka, J. (2006). Automatic online subtitling of the Czech parliament meetings. In *Proceedings of TSD 2006, Lecture Notes in Artificial Intelligence*, vol. 4188 (pp. 501–508). Berlin – Heidelberg: Springer.
- Romportl, J. (2010). On the objectivity of prosodic phrases. *The Phonetician*, this volume.
- Tihelka, D. (2005). *The Unit Selection Approach in Czech TTS Synthesis*. Ph.D. dissertation, Department of Cybernetics, University of West Bohemia, Pilsen.
- Tihelka, D. & Romportl, J. (2009). Exploring automatic similarity measures for unit selection tuning. In *Proceedings of Interspeech* (pp. 736–739). Brighton, UK.
- Zajíc, Z., Machlica, L., Vaněk, J., Padrta, A., Radová, V. (2008). An expert system in speaker verification task. In *Proceedings of Interspeech* (pp. 355–358). Brisbane, Australia.
- Zovato, E. & Romportl, J. (2008). Speech synthesis and emotions: a compromise between flexibility and believability. In *Proceedings of the Fourth International Workshop on Human-Computer Conversation*. Bellagio, Italy.

Jan Romportl & Josef Psutka
Department of Cybernetics
Faculty of Applied Sciences
University of West Bohemia
rompi@kky.zcu.cz, psutka@kky.zcu.cz

CONFERENCE REPORTS

The Sixth Foreign Language Phonetics Teaching Conference 8–10 May 2006, Mikorzyn, Poland

The sixth *Foreign Language Phonetics Teaching* (“Dydaktyka fonetyki języka obcego w Polsce”) conference took place in Mikorzyn, May 8–10, 2006. It is an annual meeting of researchers and teachers of phonetics from all over Poland, organized by Neophilology Departments of two Vocational Schools: those in Konin and in Płock. The principal aim of these conferences, which started in 2001, is bringing together foreign pronunciation specialists for intensive discussion of the most fundamental, intriguing and practical issues having to do with phonetics teaching and learning in the setting of Polish schools and universities. Some of the most frequently reoccurring topics in the six years of the conference were:

- theory versus practice in pronunciation teaching and learning
- practical versus theoretical phonetics in Polish universities
- methods and techniques in teaching pronunciation
- pronunciation and phonetics textbooks and courses
- pronunciation teaching tools and aids
- ICT in teaching and learning pronunciation
- standards and models of phonetic correctness
- syllabi and curricula in teaching pronunciation
- pronunciation testing
- the role of transcription in pronunciation teaching and learning

Programmes, reports, abstracts, selected presentations and photo galleries from all six conferences may be found at the following internet address: <http://elex.amu.edu.pl/~sobkow/>. The six proceedings volumes, containing about 130 papers altogether, can be purchased directly from the two respective schools.

The sixth conference in Mikorzyn near Konin attracted forty participants and covered a wide spectrum of phonetic interests: from learner strategies (Mirosław Pawlak), through criticism of *Lingua Franca Core* (LFC, see Jennifer Jenkins’s writings; Agnieszka Bryła), a review of available phonetics teaching materials (Jolanta Szypra-Kozłowska) to discussion of rhythm (Beata Grzeszczakowska-Pawlikowska) and ‘teacher-talk’ (Ewa Waniek-Klimczak). The full programme of the conference is available at <http://elex.amu.edu.pl/~sobkow/MikoProg.htm>. The preferred language of the conference was Polish, but some presentations were delivered in English. As a bonus to readers of this volume, professor Szypra-Kozłowska appended three reviews of recently published phonetics textbooks:

- Hancock, M. *English Pronunciation in Use*. 2003. Cambridge University Press.
- Ashby, P. *Speech Sounds*. 2005. London & New York: Routledge.
- Ashby, M. & Maidment, J. *Introducing Phonetic Science*. 2005. Cambridge: Cambridge University Press.

According to the reviewer of the proceedings volume, professor Katarzyna Dziubalska-Kołodziej (my translation – WS), “most of the submitted papers are practically oriented phonetics teaching presentations [...] There are also academically grounded papers, using a more empirical methodology, as well as review papers. The best achievements in each of the three categories are:

1. Radosław Świąciński, in his “Teaching English articulatory setting features to Polish students of English – a study of phonation”, shows that by using appropriate techniques and exercises, one can force change even in such a seemingly poorly consciously controlled phonetic parameter as the setting of the articulatory base. The experimental group achieved a statistically significant improvement over controls in mimicking English phonatory habits (e.g., vocal fold tenseness).

2. Włodzimierz Sobkowiak & Wiesława Ferlacka (“Calibrating the Phonetic Difficulty Index”) attempt an empirical ‘calibration’ of the Phonetic Difficulty Index (PDI) proposed by Sobkowiak in his writings. It turns out to be inadequate on the lower levels of English proficiency. Briefly: errors of pronunciation predicted by PDI occur rarely, while many of the actually attested errors are not accounted for by the index. The value of this paper is in the stimulus it provides for further work on this important and neglected aspect of pronunciation teaching.

3. Mirosław Pawlak’s “On the use of pronunciation learning strategies by Polish foreign language learners” is a brilliant and methodologically mature questionnaire study conducted on a large sample of subjects and based on well-selected literature. The main theme is the application of appropriate pronunciation learning strategies, i.e., the extension into phonetics of the most hotly debated field of reflection in contemporary foreign language teaching and learning theory.”

The conference programme also included three unpublished workshops:

- Włodzimierz Sobkowiak & Aleksandra Siekierska-Wojnowska: “PDI for everybody!”
- Wiktor Gonet: “Check your pronunciation”
- Geoffrey Schwartz: “Can you always trust your ears?”

The first concerned the beta version of the PDI user interface (see above), allowing interactive querying of the PDI-enabled Phonetic Access Dictionary. The other two workshops demonstrated how to use freely available acoustic tools to assist the EFL pronunciation learner in perceiving and producing subtle phonetic contrasts (e.g., degree of aspiration, vowel quality and length, voicing).

Włodzimierz Sobkowiak
School of English
Adam Mickiewicz University, Poznań, Poland
sobkow@konin.edu.pl

The 14th Manchester Phonology Meeting **25–27 May 2006, University of Manchester, United Kingdom**

The Manchester Phonology Meeting (generally known as the ‘mfm’) has been a key phonology conference in the UK since 1993. The conference has always been organised by phonologists from various universities and this year, the organising committee consisted of Patrick Honeybone (University of Edinburgh), convenor and main organiser, Ricardo Bermúdez-Otero (University of Manchester), Wiebke Brockhaus-Grand (University of Manchester), Philip Carr (Université de Montpellier – Paul Valéry/ERSS, Toulouse – Le Mirail) and Jacques Durand (Université de Toulouse – Le Mirail), helped by an international advisory board. The conference is supported financially by the Linguistics Association of Great Britain and the British Academy.

The conference website (<http://www.englang.ed.ac.uk/mfm/14mfm.html>) indicates that the aim of the mfm is to give anyone who is interested in phonology an opportunity to submit an abstract on anything related to phonology, in any theoretical framework. Jacques Durand, one of the mfm’s founders, writes on the ‘mfm homepage’ (www.englang.ed.ac.uk/mfm/mfm.html) that this open vision has been present since the mfm was founded. The first organisers recognised a need to pull phonology out of its isolation and, initially, provide an opportunity for phonologists, phoneticians and sociolinguists to meet, discuss issues and contribute to each other’s fields. Later, the need to involve fields, such as psycholinguistics and neurolinguistics, was also recognised. The conference serves as a location to discuss the place of phonology in a theory of language and its interaction with the other ‘core linguistic fields’ – morphology, syntax, semantics and pragmatics. Jacques Durand points out that the variety of topics and issues discussed has “...made the mfm different from many other phonology meetings which have limited themselves to technical issues within phonology and contributed to the splendid isolation of our field”.

The mfm also boasts a great social programme which contributes to the informal atmosphere of the conference and facilitates discussion and debate. The day usually ends with a visit to the local pub, the Whitworth. Furthermore, two dinners are organised: On the first night of the conference, a typical British tradition is honoured with a dinner in an Indian restaurant on Manchester’s famous “curry mile” close to Hulme Hall. The conference dinner on Friday consists of an excellent Chinese buffet in a restaurant in Manchester’s vibrant city centre.

The fact that the mfm is special was proved again this year by the 111 participants, representing many different universities from five continents. The conference was divided in a series of general talks in parallel sessions (46 in total), two poster sessions (with 31 posters in total) and a special session. In addition, there was a book display offered by publishers such as Blackwell, OUP, and John Benjamins. The books were the centre of a very popular book auction on the last day of the conference.

In the tradition of the mfm, many different issues were addressed, data from many different languages were presented and a wide array of phonological frameworks was discussed. Unfortunately due to limited space, only very few of the excellent papers presented at the conference can be mentioned in some detail here.

Adam Albright (MIT: “Why eatees are not E.T.’s: Blocking of aspiration by output-output constraints”) argued for an Output-Output faithfulness constraint governing the failure of aspiration in English before stressed suffixes, such as *-ée* as in *esca*[p]ée. He pointed out that word-final stops do not aspirate before stressed vowels, as in *soa*[k] úp and failure of aspiration in this context is usually attributed to (i) the fact that English does not resyllabify across certain morphological boundaries, and (ii) aspiration applies only within the syllable.

However, he showed that certain word-internal stressed vowels also fail to trigger aspiration, which occurs when certain stressed suffixes, like *-ée*, are attached to verbs ending in voiceless stops, as in the *escapee* example above. He further pointed out that this ending is productive, as can be seen from nonce forms such as *sò*a[k]ée. In addition, acoustic analysis confirms that the VOT value of stops before *-ée* is short, equivalent to stops before unstressed suffixes such as *-er*, and that the stops have a long closure duration equivalent to word-final stops. This he argued is evidence for the fact that suffixes like *-ée* truly fail to trigger aspiration. He argued that positing that *-ée* blocks resyllabification like word-boundaries do is inadequate for a number of morphological and phonological reasons (lack of morphological evidence and failure of flapping in these contexts). He argued that, rather than in terms of resyllabification, the facts are better explained in terms of an O-O faithfulness constraint which demands that stops before productively derived boundaries must closely resemble their phonetic realisation in isolated verb forms. This then explains the difference between a suffix like *-ée*, which is productive and does not cause aspiration, and a suffix like *-éer* (as in *volun*[t^h]éer), which is unproductive and therefore triggers aspiration.

Jill Beckmann, Michael Jessen and Catherine Ringen (University of Iowa & Bundeskriminalamt, Germany: “Phonetic variation and phonological theory: German Fricative Voicing”) presented experimental data in order to test the different predictions of two phonological analyses within OT of German fricative voicing: coda devoicing (an obstruent is devoiced in coda position), and positional faithfulness (an obstruent is devoiced when it is not in a pre-sonorant position). Their aim was to provide a full positional faithfulness account of the German fricative data and to show that only this account is consistent with the variation found in the data. They took the theoretical position that laryngeal contrast in German stops is one of [spread] versus no laryngeal specification. Therefore, syllable final devoicing of *stops* does not exist in German in that all stops are voiceless in the language unless they are variably (passively) voiced between sonorants. However, German *fricatives* do contrast for voice in both initial and medial position, this may suggest that coda voicing does take place in German fricatives. Both theories can deal with a word like *verlo*[s]te ‘raffle 1 SG/3SG PAST’ from *verlo*[z]en ‘raffle Inf’; coda devoicing predicts devoicing because /z/ is in a coda position and positional faithfulness predicts devoicing because /z/ is not followed by a sonorant.

However, when a sonorant follows, as in a word like *gruslig* ‘spooky’ or *fasrig* ‘fibrous’, the two accounts differ crucially. When [z] is produced there is evidence that there is no coda devoicing since the syllabification of the medial clusters is [z.l], [z.r] due to the fact that German does not allow onsets like *[zl] and *[zr]. However, the positional faithfulness account does predict the voiced pronunciation. Experimental data recorded from 36 native speakers of German points out that in these words, the fricative was

sometimes voiced and sometimes voiceless. This variation, they argue, can only be explained in the positional faithfulness analysis as variable (phonetic) failure to achieve voicing in segments in which voicing is difficult. The coda devoicing analysis does not offer a comparable explanation.

Norval Smith & Bert Botma (ACLC & LUC: “Denasalisation in Delta Yokuts”) presented data from Delta Yokuts varieties showing denasalisation of nasal consonants. Denasalisation takes place when a nasal consonant changes into an oral consonant. This process adds a previously not existing voiced series to the Delta Yokuts dialects’ stop systems.

However, stops which occur in alternation with nasals have been claimed by many authors not to be real obstruents but are argued to be sonorants phonologically. Smith & Botma discussed the available evidence on this issue in terms of a Dependency Phonology approach and drew the following conclusions: (i) in Delta Yokuts, there is a phonetic gap at the voiced stop position. Denasalisation fills this gap by extension of allophones to cover this gap. (ii) denasalisation is often blocked if there is no available gap at the voiced stop position. This can be attributed to the general tendency of languages to preserve underlying segmental contrasts. (iii) In some languages, there are good grounds to assume that denasalised stops pattern as obstruents. Denasalisation would then involve the transfer of nasals to obstruent stops to fill a phonological gap.

James M. Scobbie & Jane Stuart-Smith (Queen Margaret University College & University of Glasgow: “Covert articulation of Scottish English /r/: now you see and hear it, now you don’t”) combined sociolinguistic research and ultrasound data to illustrate the existence of weakly rhotic and derhoticised speech in broadly spoken vernacular Scots, in which coda /r/ sounds weaker than onset /r/ or is completely deleted. In their paper, they use acoustic analysis to support finely transcribed impressionistic data and showed that monophthongal and diphthongal vowels are found rather than rhotic approximants with lowered F3. Scobbie & Stuart-Smith argued that these facts suggest a diachronic change towards non-rhoticity. Additionally, difficult synchronic analytic questions about the relationship between phonology and phonetics are raised.

Then they presented new ultrasound data to exemplify some of the Scottish variation. The question of whether /r/ should be present in a phonological analysis of these facts is approached by comparing different frameworks. All these offer certain advantages, in order to understand gradient processes which eventually lead to categorical phonologisation and phonological change. The fact that a speaker may cognitively commence a process, rather than it being only an inter-speaker inter-generational change from a phonetic to a phonological representation, is suggested by the fact that in some cases a sound cannot be heard, i.e., there is no acoustic output, but it can be seen on the ultrasound, i.e., the speaker is making the underlying articulatory movements.

This year’s special session was dedicated to the memory of Peter Ladefoged and focused on “Fieldwork and phonological theory”. A few papers in the general session had already referred to this topic: Heriberto Avelino (University of California, Berkeley: “The phonetic structures of endangered Mexican languages project: consequences of fieldwork phonetics for phonological theory”), and Erich R. Round (Yale University: “The phonologist and the design of documentary fieldwork: assuming a role in data production from the outset”).

The special session was divided into two blocks of 1.5 hours followed by some time for discussion and debate.

The first block consisted of a talks by Daniel L. Everett (University of Manchester: “On the philosophy of field research and the relationship of field research to phonological theory”) and Larry Hyman (University of California, Berkeley: “Phonological Theory and field work: is the gap widening?”).

Everett introduced his talk by giving a definition of fieldwork and discussing the history of field research in terms of colonialism, and descriptivism versus structuralism, and its philosophy in terms of what he calls “Homeopathic bias and normal science” and “Raging against the truth”. He then went on to discuss methodology and theoretical issues by taking a look at standard segmental phonology methodology and its inadequacies. He discussed the importance of minimal pairs in this framework and argued that when these cannot be found the linguist should look for contrast in analogues environments which can be exemplified by pairs like *spit* and *Pete*. He then provided an example from Pirahã tone analysis to illustrate the point that a theory which only incorporates minimal pairs cannot always adequately deal with the facts of languages. A second type of methodology discussed is ethnophonological methodology. He shows that, for instance, /s/ is missing from women’s speech in Pirahã. Ultimately, he described Postal’s Maxims which state that field workers should learn the language they investigate as well as possible, and an attempt should be made to “...formulate an explicit account of the rules which generate the full syntactic structures of its sentences”. Everett then discussed a wide array of philosophical approaches to field work: the entrance of American Indian philosophy into American European philosophy, the Pragmaticist turn, the Anthropological view, American Descriptivism and Rara, American Structuralism and Rara, The Chomskian view: the Truth of Language, Contingency and Chance, Individual Phenomena versus the whole, The importance of the individual, Utility, Coherence, Radical Empiricism, tolerating inconsistencies, and finally Sapirian Descriptivism. Everett concluded that Postal’s maxims should form the methodology of field work and the philosophy adopted towards it should be Pragmatism.

Hyman’s paper addressed two questions with regard to the relationship between phonological theory and fieldwork: (i) what counts as (phonological) “field work”? (ii) what counts as (phonological) “theory”? As an additional goal, it questioned what phonological field work and theory can do for each other. In order to answer these questions he addressed a number of other questions relating to the goals and importance of field work and phonological theory, and who should do it. He then went on to argue that it is very difficult to separate phonological typology and phonological theory and addressed the question of how phonological theory and (descriptive) field work are currently contributing to each other and how it looks in terms of their future interactions. He presented five frameworks which provided effective tools to look at certain aspects of phonology, creating a descriptive “boom”: Structuralist Phonology, Classic Generative Phonology, Non-linear Phonology, Lexical Phonology, which all provided concepts and tools that have informed and facilitated the work of field phonologists, and Optimality Theory. He set OT apart from the others by questioning whether it is desirable, or even possible, to do “non-theoretical” language specific work in this framework, and what the concepts or mechanisms useful to field work are in OT. He presented “floating tones” data

from various languages to discuss this point in detail. His concluding questions about phonological theory and field work are: (i) does phonological theory provide the tools needed by a field worker?, and (ii) what would a theoretical phonologist like from field work? He argued that phonological theory should provide the field worker with formal models which (a) express insights and solve problems and (b) discover new insights and identify new problems. Field work should provide phonologists with analyses which are (a) rigorous and comprehensive and (b) rich, insightful and interesting. He concluded by arguing that phonological theory and field work are bound together by a shared concern for typology and analysis and although research agendas differ, they cannot be objectively ranked.

The second block of the special session consisted of a talk by Karen Rice (University of Toronto: “Free variation in Slave (Northern Athabaskan): relating field work and phonology”). Rice pointed out that field work these days requires a set of ethical responsibilities to the community, including working with the community in decision making and respecting the cultures, knowledge, and values of the people whose language is investigated. Language is a particularly sensitive area of study. She points out that the Slave language, an Athabaskan language of northern Canada, displays a lot of variation which is important in the community: variation marks both individuals and geographic areas and is found within the speech of individuals. In the first part of the talk, Rice examined the variation in phonetic output both from a theoretical and language-internal perspective. Theoretically, she argued that a model of phonology that builds inventories monotonically through a set of choices in a fixed segment structure offers insight into parts of the phonology where variability, both cross-linguistically and language internally, is potentially available. She further argued that the actual implementation of variation is due to external factors, including phonetics and social factors. She illustrated these points with data from two cases of variation in Slave, variation between a palatal glide and a voiced alveopalatal fricative, and variation between a nasal, a prenasalised stop, and a voiced oral stop. She then examined the implications of variations for choices of orthography and argued that suppressing free variation in spelling leads to extra difficulties for the writer and causes loss of information about individual speakers, whereas this information is deeply valued by the community. She concluded that linguists have a responsibility, in working on a language for theoretical purposes, to also be responsible to the particular community in which they are working, and give to that community in ways that are appropriate.

The special session ended with a remembrance of Peter Ladefoged by Jacques Durand and Daniel L. Everett (CNRS, Toulouse 2 & University of Manchester).

Some of the excellent posters presented at the mfm included “Phonology between home and field research” by Nabila Louriz (Hassan II University, Casablanca), “Statistics is not enough for language acquisition” by Naomi Yamaguchi (LPP, CNRS-Paris III), “Faithfulness and identity in Luganda reduplication” by Francis Katamba (Lancaster University), “Phonetic cues for syllable structure? Evidences from labiovelars in Tuscany” by Nadia Nocchi (Phonogrammarchiv, Zürich), “The perception of L2 stress” by Heidi Altmann (University of Delaware), “Towards a ‘quantal’ definition of nasal vowels, on the basis of psychological and acoustic evidence” by Angélique Amelot (LPP, UMR 7018 CNRS-Paris 3), “Neural correlates to a three way contrast of duration in speech and non-speech” by Herberto Avelino & Anna Shestakova (University of California, Berkeley,

University of Helsinki), “MWUs and the prosodic hierarchy: a case study” by Nicholas Ballier & Susan Maroux (University of Paris 13, University of Limoges), “A three-way comparison in perceptual development: monolingual children vs. adult L2 learners” by Ivana Brasileiro & Paola Escudero (Utrecht University, University of Amsterdam), “The influence of phonemic vowel-length on the voicing effect” by Rebeka Campos-Astorkiza (University of Southern California), “A template for Turkish” by Ann Denwood (Bosphorus University), “Root-and-No Pattern Morphology” by Dafna Graf (LUCL, Leiden University), “Incongruent speech data: its potential in phonological fieldwork” by Michael Ingleby & Azra N. Ali (University of Huddersfield), “A feature-driven loanword adaptation of English and French plosives into Korean: a case of symbiosis between perception and grammar in loanword phonology” by Hyunsoon Kim (Hongik University, Seoul), “Nasal consonants in Tupi and Jê languages” by Luciana Storto & Didier Demolin (Universidade de São Paulo, Université libre de Bruxelles), and “Coping with ‘non-ideal’ utterances: why speakers don’t have to be perfect” by Frank Zimmerer & Henning Reetz (University of Konstanz, Johann Wolfgang Goethe University).

Marleen Spaargaren
University of Edinburgh, Great Britain
s0348141@sms.ed.ac.uk

XIth International Conference SPECOM’2006 25–29 June 2006, St. Petersburg (Russia)

The 11th International Conference “Speech and Computer” was held in St. Petersburg, Russia on June 25–29, 2006. SPECOM was established in 1996 by the St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS) together with the State Pedagogical University of Russia. Since then, SPECOM has also been held in Moscow by Moscow State Linguistic University (MSLU), in Cluj-Napoca (Romania) by the local centre of the Research Institute for Computer Technology, and the 10th jubilee SPECOM’2005 was hosted by the University of Patras, Greece.

SPECOM’2006 was organized by SPIIRAS, with financial support of the SIMILAR Network of Excellence, the INTAS International Association, the Administration of St. Petersburg, the Russian Foundation for Basic Research, the ICA International Commission for Acoustics, and the U.S. Army Research Laboratory – European Research Office. Supporting organizations were ISCA, EURASIP, ELSNET, as well as the service agency Monomax Ltd.

The final technical programme included 109 papers by 262 authors from 33 countries presented in 11 oral, poster and demonstration sessions. All papers had been subjected to a thorough review by members of the Scientific Committee, composed of 32 distinguished reviewers. Paper acceptance ratio was 74.2% for the regular sessions.

The development of natural means of human-computer interaction is currently becoming one of the main trends in informatics. The continuously growing performance of computers and network technologies is not yet fully exploited owing to the “unnatural” character of the human-computer dialogue. Insufficient solutions to this problem inhibit

the development of various applied systems in telecommunications, medicine, edutainment and everyday life, since all modern technologies and network services use automated devices for information control and processing. Every year, the SPECOM conference aims to provide all professionals dealing with human-computer interaction with new scientific and educational knowledge. This forum brings together colleagues from all over the world to exchange knowledge related to various aspects of human-computer interaction. The conference venue and dates were selected so that the participants can possibly be exposed to St. Petersburg's unique and wonderful phenomenon known as the White Nights, for our city is the world's only metropolis where such a phenomenon occurs every summer.

Before the conference, the chairman of SPECOM'2006, director of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), professor Rafael Yusupov, had emphasized the importance of the conference. He stated that the global aim of this event was to discuss the state-of-the-art in speech technologies and attract different specialists (engineers, mathematicians, philologists, doctors, teachers, psychologists) to the problem of natural interaction between humans and computers. There is a need for fusion between diverse areas of science to create systems which are able to understand and generate speech, handwritten text, natural gestures, head and body movements.

On the first conference day, June 25, 2006, two tutorials devoted to problems of recognition and synthesis of speech were held. Professor Lawrence Rabiner (USA) gave the tutorial "Challenges in speech recognition and natural language understanding", and professor Thierry Dutoit (Belgium) gave the tutorial "Corpus-based speech synthesis".

The tutorial day was opened by professor Thierry Dutoit who presented a survey of modern methods and systems for automatic speech synthesis. Owing to new corpus-based methods and the application of automatic unit selection techniques, modern speech synthesis systems have reached a quality which experts could not predict even ten years ago. Usage of TTS technologies has allowed improved functions of mobile devices and provided new capabilities for telecommunication services. The main scientific problems in this area are now concerned with prosodic analysis, intelligibility, compression of speech corpora for application of such systems in mobile phones, as well as with the development of flexible ways for voice cloning and creation of new voices.

In the next tutorial, professor Lawrence Rabiner presented the state-of-the-art of speech recognition as observed in commercial systems with English ASR. Natural language understanding and automatic speech recognition techniques have reached acceptable quality and now can be applied to dictation tasks, dialing phone numbers, automated enquiry services and call-centres. But in spite of the progress, the main challenge in this area is the robustness of ASR systems in real-life conditions. A system must work reliably in any environment, for any speaker and with any application. In the tutorial, professor Lawrence Rabiner presented all levels of speech processing involved in full speech dialogue with a computer: automatic speech recognition, natural language processing, dialogue management, generation of an answer, and speech synthesis.

During the opening of the conference on June 26, the deputy director of SPIIRAS, professor Boris Sokolov, gave the welcome speech and mentioned the ultimate importance of research in ASR and NLP domains, as well as in human-computer interaction as a

whole. The importance of new directions in science and technology and that of multimodal interfaces in the world was emphasized.

Then professor Rajmund Piotrowski (Russia), professor Lawrence Rabiner (USA) and professor Christian Wellekens (France) gave their welcome speeches.

Professor Rajmund Piotrowski welcomed the participants of the conference on behalf of all linguists in St. Petersburg. He reminded the audience that the first SPECOM dedicated to man-machine speech interaction was also held in St. Petersburg, and chaired by Professor Yuri Kosarev, 10 years ago. Professor Piotrowski noted that this forum has survived a difficult journey. One time, applied linguistics was close to death in Russia but owing to development of information technologies and young specialists, new horizons were reached.

Professor Lawrence Rabiner noted that speech technologies have also come a long way, but the problem of ASR, in spite of unquestionable successes, cannot be considered as solved. He continued describing the tasks involved in ASR, which he presented during the tutorial, and mentioned an evaluation technology which consists of four stages. In the first stage, the basics of acoustics, phonetics and linguistics are considered. In the second stage, the speech representation and fundamentals of speech perception are studied. The third stage includes the construction of the algorithms for tone evaluation, formants, voice, and silence, as well as other parameters of the speech signal. During the fourth stage, the integration of all methods in one system is performed. He also referred to problems of speech coding, synthesis and verification.

Professor Christian Wellekens pointed out the erroneous belief that human-computer speech interaction is identical to human-human interaction. He noted that since the 1970s, speech processing has been characterized by rapid evolution, and emphasized the importance of a the joint work of engineers, phoneticians and philologists in modelling the processes of ASR. He then mentioned the problem of robustness in speech recognition and understanding. In conclusion, professor Christian Wellekens described the activities of the ISCA association.

After the opening ceremony, professor Christian Wellekens (France) gave the keynote lecture. Further lectures were given by professor Rüdiger Hoffmann (Germany) and professor Christoph Draxler (Germany) on June 27 and 28.

Professor Christian Wellekens presented a survey lecture entitled “Impact of variabilities on speech recognition”, in which he mentioned the progress of ASR, but also described the technological barriers preventing us from creating robust systems. This issue is connected first of all with high sensitivity of digital speech processing to diverse external noises, as well as with approaches of grammar rules and semantic representations which are not well-developed. Special attention should be paid to research on natural variability in speech. For instance, the accuracy of a system decreases sharply when a speaker has a foreign accent. There are many factors which influence the quality and style of speaking: environmental, regional, socio-linguistic and personal factors. All this variability should be taken into account while developing robust ASR systems.

Professor Rüdiger Hoffmann presented a historical survey under the title “Speech synthesis on the way to embedded systems”. At present, TTS systems fall into two groups: (1) systems able to use “unlimited” computer resources – server applications, and (2)

speech synthesis for mobile devices where there are strict limitations in memory volume and computer performance. The major part of TTS market is oriented towards mobile devices, therefore developers have to provide natural-sounding synthetic speech with minimum requirements in computational resources.

In his lecture “Web-based speech data collection and annotation”, professor Christoph Draxler discussed the necessity of large speech corpora of continuous and spontaneous speech for ASR. He suggested that the efforts of scientists and commercial companies be united to collect speech databases via the Internet. He also presented web services intended for speech recording and processing. Such an approach was successfully used in Germany, and is scheduled to increase the number of participants collecting speech corpora for many languages, which is required for effective training of ASR systems.

The multimodal interfaces day was organized on June 28, 2006. It included several oral presentations on natural signal processing, as well as the round table “SIMILAR NoE brainstorming on multimodal semantic fusion”, dedicated to the problems of semantic information integration. Professor Benoît Macq (Belgium) opened the round table and briefly presented main approaches for semantic information fusion and fission which are used in the development of multimodal interfaces. In the framework of SIMILAR NoE, the OpenInterface platform has been created, which will allow scientists to carry out joint research and integrate natural modalities by researchers from different countries. Professor Niels Ole Bernsen (Denmark) gave a lecture on the integration of speech and gestures for control and conversation with virtual characters. By means of the NICE system, where a virtual H. C. Andersen talks about his life, fairytales and museum in Odense, main features of a multi-modal interface were demonstrated, along with the problems of synchronization and fusion of gestures and speech commands.

The official representative of the INTAS association, Dr. Yuri Melnikov, gave a speech on the priority directions of the Association’s activities and presented some results of completed and ongoing research projects in the area of information technologies. The role and means of development of INTAS in the future EU FP7 programme were discussed. The representative of RFBR, professor Nelli Didenko, presented a talk on grants for fundamental research in St. Petersburg, as well as cooperation between scientific and international funding organizations.

At the completion of the conference, the scientific committee acknowledged the success of the 11th International Conference “Speech and Computer” SPECOM’2006 in St. Petersburg. During the four days, the participants could share their experience in development and application of speech and multimodal technologies in industry, telecommunications, medicine, culture and education for solving current problems of human-computer interaction, including the attraction of young researchers to the fundamental tasks of computer science. More details about the conference, scientific committee and special events are available at <http://www.specom.nw.ru>. The 12th International Conference “Speech and Computer” SPECOM’2007 will be organized by Moscow State Linguistic University and SPIIRAS.

Andrey Ronzhin
St. Petersburg Institute for Informatics and Automation
Russian Academy of Sciences
ronzhin@ias.spb.su

Beszéd kutatás 2007
(Speech Research 2007 National Conference)
15–16 November 2007, Budapest (Hungary)

Beszéd kutatás (Speech Research) conference takes place every second year in Budapest, at the Research Institute for Linguistics of the Hungarian Academy of Sciences. The organizers of the conference are professor Mária Gósy (RIL, HAS and ELTE Department of Phonetics) and her colleagues both at the Institute and at the University. A large number of researchers in several disciplines related to speech science from all over Hungary present their current research at each Beszéd kutatás conference. Short 10-minute presentations on hot topics in phonetics, psycholinguistics, neurolinguistics, speech technology, clinical linguistics, forensic phonetics, and other linguistic and psychological disciplines are always welcome. The abstracts submitted for evaluation are peer-reviewed by the organizing committee. Key talks are given on both days of the conference.

This year, the subtitle of the conference was “Phonetics and Psycholinguistics”, indicating the importance of the cooperation between various sub-fields of speech science. Professor István Kenesei, director of the Linguistics Institute, greeted the participants of the conference and gave a short overview of the previous conferences.

The first day plenary talk of Beszéd kutatás 2007 was given by Professor József Hámori, vice-president of the Hungarian Academy of Sciences. The topic of his talk was the relationship of brain development and language. The evolution of the human language requires a general usage of right-handedness, on the one hand, and the morphological changes of the brain, on the other hand. Two functional copies of the FOXP2 gene seem to be also required for the acquisition of spoken language.

Three books were presented during the first session. Anna Adamikné Jászó presented the recently published book about speech perception deficits, edited by Mária Gósy. The book contains the results of recent investigations on (i) the speech perception and comprehension processes of typically developing children and (ii) decoding processes of children with various disorders. The book, written by Judit Navracics, discusses the working of the bilingual speakers’ mental lexicon, based on several years of research. Alexandra Markó presented a book about Hungarian consonants and their coarticulation. The author is Professor Gábor Olaszy, who had also a presentation about the database connected to the book.

The subsequent sessions contained presentations of recent investigations. There were 46 talks, which could be sorted into 11 major topics (Fig. 1).

Sixteen presentations were given on segmental or suprasegmental questions of Hungarian. The presenters investigated articulatory aspects of the Hungarian vowel system (Katalin Mády), acoustic characteristics of /v/-realizations (Zsuzsanna Bárkányi & Zoltán Kiss; Tamás Bóhm & Gábor Olaszy), and coarticulatory aspects of sound combinations (Gábor Olaszy; Tekla Etelka Grácsi). Gábor Kiss and Mika Waseda presented their comparative study of Hungarian and Japanese vowel systems. Kálmán Abari and Gábor Olaszy demonstrated their phonetic database which offers the possibility of analyzing the temporal structure of words on the internet.

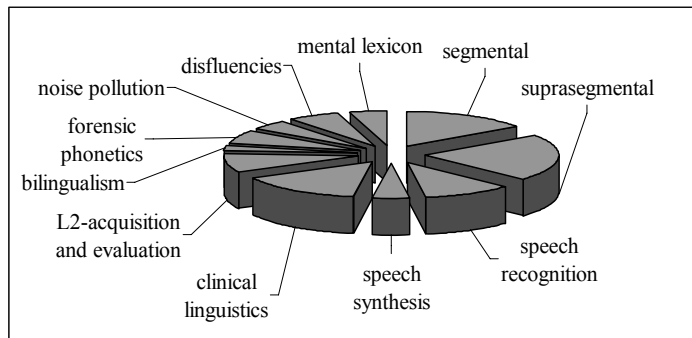


Figure 1: Major topics at the Beszédkutatás 2007 conference

Presentations about suprasegmental problems concerned both traditional and newer issues in Hungarian phonetics. Production and perception of accent in spontaneous speech was examined by Angéla Imre and Judit Bóna. Melody and stress patterns of reproductive speech were analyzed by János Nagy. Alexandra Markó analyzed question realizations in spontaneous speech. Speech tempo was discussed from a diachronic perspective, comparing the sound of a film shot in the 1940s with its recent remake (Attila Mártonfi). Perceptual characteristics of speech in various tempos were examined by Krisztina Menyhárt. Speech tempo in teenagers' speech production and dysfluencies were analyzed by Mária Laczkó. Judit Kozma presented a comparative study of pitch movements in Hungarian and German reading. Sarolta Bata presented her results on segmental and suprasegmental factors in the temporal organization of spontaneous speech. The talk given by Viola Váradi dealt with perceptual patterns of spontaneous speech utterances.

Seven talks were given on speech technology. Their topics concerned emotions in speech: namely aspects of their recognition (Zoltán Tüske, Márta Simon, Péter Mihajlik & Tibor Fegyó; Szabolcs Levente Tóth, Dávid Sztahó & Klára Vicsi) and synthesis (Csaba Zainkó and Márk Fék). György Szaszák and Klára Vicsi presented their speech recognition system using prosodic information as a performance-enhancing factor; Zsolt Németh, György Szaszák and Klára Vicsi talked about modality detection. Anne Tamm, Kálmán Abari, Gábor Olaszky and Kata Gábor raised several methodological questions concerning syntax-oriented implementation of word stress in speech synthesis.

Speech production in aphasia (Zoltán Bánréti, Éva Mészáros and Katalin Szentkúti-Kiss), consequences of severe hearing-impairment on language acquisition (Mónika Bombolya), Alzheimer's disease (Laura Szél), speech properties after using marijuana (Zsuzsa Kaló) and language abilities of children with learning difficulties (Mónika Macher) were investigated by authors dealing with clinical linguistics. Talks on L2 acquisition concerned pronunciation difficulties (Ágnes Kuna), spontaneous speech production of non-native speakers (Erzsébet Balogh), and evaluation methods and their reliability (Judit Kormos and Mariann Dénes). Perceptual processes connected with children's speech are an important topic, both on a theoretical and practical level. Orsolya Simon investigated these processes in the children's L1 and L2 productions. The talk given by Judit Navracsics on bilingualism analyzed the effects of the speech partner on the planning and code switching in early and late bilingual people.

There were three presentations on forensic phonetics about (i) the usability of pitch movements and accentuation in speaker identification (Zoltán Tatár, Zoltán Vargha & Attila Fejes), (ii) the methodology of speaker identification using various programs (Zoltán Vargha) and (iii) noise reduction (Attila Fejes).

Two papers concerned the effects of environmental noise on speech. Mária Gósy analyzed the occurrence of dysfluencies and the suprasegmental consequences of noise in speech production (Lombard effect); the effects of different types of noise on speech were analyzed by Dorottya Gyarmathy. Dysfluencies were also investigated in the speech of simultaneous interpreters by Mónika Kusztor and Mária Bakti. An analysis of dysfluencies with respect to their word class, based on the Hungarian speech error corpus, was carried out by Viktória Horváth. Syntactic aspects of dysfluencies in Hungarian were investigated by Tamás Bíró.

The structure and function of the mental lexicon is one of the main topics in psycholinguistics. Hidden processes, such as the ways of searching in the mental lexicon (Ákos Gocsál & Ágnes Huszár), and the lexical access in word comprehension (Orsolya Csiszár) were investigated through various experiments.

The next conference will be held in 2009 – the main topic will be spontaneous speech. The organizers welcome researchers both from Hungary and from abroad, who investigate the following topics: phonological processes in spontaneous speech, prosody and dysfluencies, turn-taking and other characteristics in discourse, and various other topics related to spontaneous speech (speaker's age, acoustic-phonetic properties, speech perception, speech technology and databases, forensic speaker identification, etc.).

Tekla Etelka Grácsi & Viktória Horváth
Research Institute for Linguistics
Hungarian Academy of Sciences
graczi@nytud.hu, horviki@nytud.hu

**The 2nd Czecho-Slovak Conference
of the International Society of Phonetic Sciences
18 January 2008, Prague, Czech Republic**

Four years after the first event of this kind, the Prague Institute of Phonetics hosted the second Czecho-Slovak Conference of the ISPhS. This one-day event brought together Czech, Slovak, French and Swiss phoneticians and experts from related scientific disciplines. Among the ten papers presented, four were by authors from the Czech Republic, four from France, one from Slovakia and one from Switzerland. Apart from its scientific content, the conference also had a festive dimension: it was meant as a celebration of the 70th birthday of Marie Dohalská, professor at the Institute of Phonetics and vice-president of the ISPhS.

In their talk, “Régularité rythmique en français spontané” (Rhythmic regularity in spontaneous French), Katarina Bartkova and Natalia Segal (France Télécom R&D) first analyzed the occurrence of regular rhythmic patterns (successive prosodic groups of the

same length) in a French speech database. In the second part of the paper, they compared two approaches of automatic segmentation into prosodic units: the first one was based on vowel duration modelling, and the second one on prosodic tree building using linguistic knowledge, including the ratio of vowel durations and f_0 slopes.

Georges Boulakia (Université Paris 7) presented a talk entitled “Mais quelle prosodie ?!” (But what kind of prosody?!), where he mentioned different communicative aspects of prosody in several languages, with interesting diachronic observations.

In the next paper, “Rôle des contraintes phonologiques sur la reconnaissance des mots: l'exemple de deux dialectes du français” (The role of phonological constraints in word recognition: example of two French dialects), Jean-Yves Dommergues (Université Paris 8) analyzed the phonological status of vowel length in Parisian and Neuchâtel French, and showed that this difference had an impact on word recognition in both dialects.

Tomáš Duběda (Charles University in Prague) presented the paper “Variantnost melodických akcentů ve čtené češtině” (Variability of pitch accents in read Czech), where he discussed the advantages and drawbacks of intonational stylization by means of discrete tones as opposed to the more traditional contour approach, and examined pre-nuclear intonational patterns in read Czech, using an annotation system based on bitonal pitch accents.

The next speaker was Eric Keller (Lausanne University, Switzerland), whose interdisciplinary paper “Le défi du fonctionnement coopératif” (The challenge of cooperative behaviour) described different aspects of cooperative interaction within a group and underlined the fact that phonetic, semantic, conversational and gestural phenomena play an important role in social interaction.

Jana Králová (Charles University in Prague) presented the paper “Varianty španělštiny a humor: několik poznámek k didaktické prezentaci variant výslovnosti španělštiny” (Varieties of Spanish and humour: some remarks on pronunciation variants and their role in language teaching). She illustrated some possibilities for the teacher to present variants of Spanish pronunciation in language teaching, using short amusing texts.

Philippe Martin (Université Paris 7) talked about “Postfixes et suffixes interrogatifs : un cas d'ambiguïté prosodique?” (Interrogative postfixes and suffixes: a case of prosodic ambiguity?). He discussed cases of dislocated phrases and their prosodic realization, depending on the modality of the sentence, its informational structure (topic vs. comment), as well as its syntax.

Slavomír Ondrejovič (Ludovít Štúr Institute of Linguistics, Bratislava, Slovakia) described current trends in Slovak pronunciation from a sociolinguistic perspective. In his paper, “K dvom neuralgickým bodom slovenskej sociofonetiky” (On two critical issues of Slovak sociophonetics), he primarily discussed the status of the so-called rhythmic law in today's Slovak, as well as that of /l/ palatalization.

The next paper, entitled “Vliv věku dětí na klasifikaci českých samohlásek” (The role of age in children's classification of Czech vowels), was presented by Jana Tučková (Czech Technical University in Prague). It described the design and realization of a corpus of healthy children's speech, between the ages of four and ten. These samples were utilized in

comparative analyses with child patients' speech during treatment. The author also showed methods of automatic vowel classification permitting speech diagnostics.

Finally, Robert Vích (Academy of Sciences of the Czech Republic) and Jan Nouza (Technical University of Liberec, Czech Republic) focused on intelligibility assessment of speech processing algorithms, by means of a specific method. The title of their paper was "Application of speech recognition and rhyme tests for assessment of Czech speech processing systems".

The conference papers (except for two) were published in a Proceedings volume, which also included two other papers, not presented at the conference:

The paper "Analyser, représenter et interpréter la prosodie" (Analyzing, representing and interpreting prosody) by Albert Di Cristo (Université de Provence, France) offered a general view of contemporary prosodic science. It discussed, among other things, the levels of intonational analysis, different approaches to intonational stylization, and the functions of prosody.

The paper "Maskovanie vokálov šumom" (Vowel masking by noise), written by Ján Sabol and Július Zimmermann (Prešov University, Slovakia), described experiments with vowel masking and its influence on communicative functions.

As mentioned above, the 2nd Czecho-Slovak Conference of the ISPhS was an opportunity to celebrate an important event: professor Marie Dohalská from the Prague Institute of Phonetics (Charles University in Prague) reached the age of 70 a couple of months before the event. Thus, let us devote the final lines of this report to her scientific career.

Marie Dohalská has been associated with the Institute of Phonetics since the 1960s and has been active in research and teaching, especially in the fields of Czech and French phonetics, speech communication, speech synthesis and elocution practice. Through her major works "Analyse spectrographique des voyelles françaises" (Spectral analysis of French vowels, 1977), "Dynamika verbální komunikace" (Dynamics of speech communication, 1991) and "Fonetika francouzštiny" (Phonetics of French, first published in 1992), she significantly contributed to three different branches of phonetic sciences, both on the theoretical and practical levels. Marie Dohalská was in charge of a number of research projects, some of which were international (e.g., the European COST 258 project). She has devoted much energy to developing international relations, mainly with French-speaking countries, and has co-directed several international Ph.D. theses. Her teaching skills were also appreciated abroad: she was invited twice to give lectures at the Université Paris 7. Among Marie Dohalská's activities on different boards and in different organizations, it is important to mention her membership in the Academic Senate of Charles University in Prague, and her position of vice-president in the ISPhS. Marie Dohalská has been awarded the "Ordre des palmes académiques" by the French Government and the Silver Medal of Charles University.

Tomáš Duběda
Institute of Translation Studies
Charles University in Prague
dubeda@ff.cuni.cz

BOOK REVIEWS

Leda Bisol and Cláudia Regina Bescancini, eds. (2008)

Contemporary Phonology in Brazil

Newcastle, UK: Cambridge Scholars Publishing (314 pp., incl. Contents [pp. v-xi], List of illustrations [pp. xii-xiii], List of tables [pp. xiv-xvi] and Editor's preface [xvii-xix]. Hard cover: ISBN 13: 978-1-84718-540-2)

Reviewed by:

Ben Hermans

Royal Dutch Academy of Arts and Sciences, Meertens Institute,
Amsterdam, Netherlands
e-mail: ben.hermans@meertens.knaw.nl

In this volume, fifteen articles have been collected by Bisol and Bescancini. They were all presented at a workshop on phonology organized at the Pontificia Universidade Católica do Rio Grande do Sul, in Porto Alegre, in April, 2007. The book is divided into five parts. The first part contains three studies on the prosodic phonology of Brazilian Portuguese (BP); the second part presents three studies on aspects of the historical phonology of BP; the third part contains three articles about the segmental phonology of BP; the fourth part consists of two papers on the acquisition of the language; the fifth part, finally, presents three studies about Brazil's indigenous languages.

Most of the studies are quite interesting, but the quality of the contributions varies quite a bit. Almost all of the studies deal with questions that play a prominent role in current discussions in theoretical phonology and phonetics. They show that modern Brazilian phonologists and phoneticians increasingly want to contribute to these discussions. There can be no doubt that the field can only benefit from this, as the current volume indicates. I should point out, though, that the volume suffers from one problem; in some of the contributions, there are quite a few mistakes in the way the English language is used. This is a point our Brazilian friends should pay more attention to.

In the rest of this review, I will go through all the studies, presenting their contents in a few sentences and, where necessary, also pointing out some of the hidden problems.

In Chapter 1, Moraes and Colamarco investigate the role various compensatory processes play in the intonational phonology of Brazilian Portuguese. Cross-linguistically, there are two basic patterns: compression and truncation. They are put into operation when a given melody is mapped onto a segmental string which is too short to realize all the tones of the melody. In compression, the tones are maintained, but considerably shortened phonetically. In truncation, on the other hand, part of the melody is not realized at all. Moraes and Colamarco show that BP is a language where compression is the dominant pattern. They show that this is true for all melodies of the language (yes-no question, declaration, exclamation, irony and continuation). Not only is their conclusion based on measurements of actually produced forms, it is also based on perceptual tests. Intriguingly, they show that sometimes yet another process applies when a segmental string is too short.

In this process, some of the tonal features of the melody are replaced by other, different features. Although this process has also been attested in other languages, they note that it still remains to be investigated under what conditions it applies in BP.

In Chapter 2, Magelhães analyses the distribution of main stress in BP. He develops a constraint system based on the framework proposed in Hyde (2001). This is a theory in which metrical constituents and the gridmarks (stressed positions) accompanying the constituents are represented in different dimensions or planes. Magelhães shows that his analysis can account for all characteristics of main stress in BP; not only can it explain quantity sensitivity, but also the trisyllabic window and irregular stress. It is an interesting feature of the analysis that the extrametricality device, so essential in all previous accounts of BP stress, is dispensed with. Instead of this device Magelhães introduces a kind of empty position on the grid. It is empty in the sense that it is not realized phonetically as stress, although it does count as a grid position. With this empty grid position, Magelhães can also derive other fundamental aspects of metrical theory, like a foot's minimal and maximal size. In this way, he is able to find a uniform explanation for three apparently unrelated phenomena: extrametricality, minimal size and maximal size. I should point out that Magelhães' article is extremely technical, and it is therefore not easy to understand in detail. Without any doubt this is a consequence of the limited space allotted to the author. Yet, the attempt to find a uniform explanation for three apparently unrelated phenomena deserves the reader's unconditional sympathy.

The authors of the third chapter, Bernadete, Abaurre, Romani and Svartman, study the interaction between secondary stress and various processes that affect the number of syllables. These are degemination, elision, diphthongization and reduction of high vowels. They show that in almost all cases the application of these processes improves the prosodic structure of a word, in the sense that after their application an alternating rhythm is created. The same is often true for reduction, though not always. Depending on the surrounding consonants, a high vowel can be deleted even if that leads to a stress clash. From a prosodic point of view, this is clearly not a desirable situation. The authors suggest that processes that are of a purely segmental nature, in the sense that surrounding segments play a decisive role, somehow take precedence over the principles of rhythmic organization.

I would like to point out that the authors seem to define prosodic optimization purely in terms of binary rhythm. Defined in this way, it is clear that (most of) the processes mentioned above tend to improve prosody. However, if foot structure were taken into account, then this would become less clear. Consider a case like <me oferecer>. In this example, diphthongization, creating *m[jo]ferecer* (where stressed syllables are underlined), applies at a rate of 76%. Since diphthongization creates a nice binary rhythm, one could indeed say that the process optimizes a word's prosodic structure. Looking at foot structure, however, this is not so obvious anymore. After the application of diphthongization, the foot structure of this example would be *m[jo](fere)(cer)*, where the brackets indicate foot structure. In this representation, the first syllable is unfooted. It must be unfooted, because BP does not have iambs and the first syllable is unstressed. In general phonologists claim that unfooted syllables are non-optimal. Now compare this with the pronunciation *m[ɪ]oferecer*, where diphthongization has not applied, and which is attested in only 12% of the tokens. The foot structure of this form would be *(m[ɪ]o)(fere)(cer)*,

where every syllable is nicely parsed into a foot. From the point of view of foot structure then, this realization would, strictly speaking, be preferable. That would mean that diphthongization would be disfavored. The results of the paper, then, seem to indicate that binary rhythm is *not* generated by foot structure, as is assumed by the great majority of the theoretical phonologists (cf. Hayes, 1995). The facts of BP seem to indicate that binary rhythm should be represented in some other way, for instance merely in terms of the grid, not in terms of foot structure. This surely is a highly significant result, one that should have been made explicit by the authors.

In Chapter 4, Massini-Cagliari studies the hiatus positions of BP and Archaic Portuguese (AP). She shows that in both stages essentially the same repair mechanisms eliminate hiatus. These are diphthongization, elision and degemination. However, the precise conditions regulating each mechanism in AP differ from those in BP. Some of those differences are rather striking. For instance, whereas in BP, elision only applies when the word-final vowel is /a/; however in AP, elision never applies when /a/ is word-final. In this respect, the two stages are each other's opposite. Another interesting difference is that prosodic factors, like stress clash, do not block the elision in AP, whereas they do so in BP. The conditions regulating diphthongization also differ. In AP, it is restricted to the pronouns *mi/ti*; in BP, it has a much broader application. Some of these differences, Massini-Cagliari argues, can be explained in terms of important differences between the grammars of the two stages. Thus, AP allowed two neighboring /a/s to be syllabified as a long vowel. AP, then, in stark contrast to BP, still allowed long vowels. Massini-Cagliari does not consider the possibility that the relative acceptability of long vowels in AP could perhaps also explain the fact that AP applies elision in a stress clash environment, whereas BP does not do so. Yet, I think it is quite possible to explain this difference between AP and BP along these lines. Consider the hiatus e+V, where V represents any vowel, underlining indicates stress, and '+' the word boundary separating the two vowels. In a language that allows long vowels, the loss of the stressed e is still recoverable, so to speak, in the lengthened vowel. Thus, on the assumption that AP lengthened its stressed vowels in elision context (a kind of Compensatory Lengthening effect licensed by the stress), it becomes perfectly understandable why at this stage Portuguese allowed elision under clash. It is a consequence of the fact that the removed vowel is recoverable and interpreted as part of the lengthened, adjacent vowel.

The study by Massini-Cagliari shows quite a few interesting differences between the two stages of Portuguese, and it remains to be seen to what extent these differences can be explained with general principles. In my opinion, Massini-Cagliari has made a significant first step in this direction.

In Chapter 5, Battisti explores the conditions under which clitic+host combinations were spelled without a break in 19th century Brazilian documents. Examples of this type of 'erroneous' spelling are <doaltar>, instead of <do altar> and <onosso> instead of <o nosso>. Some of Battisti's results are quite interesting. They show unambiguously that important crosslinguistic tendencies govern the spelling of clitics in these 19th century documents. Thus, she shows that there is a significant difference between clitics followed by vowel-initial hosts, on the one hand, and clitics followed by consonant-initial hosts, on the other. The latter, but not the former, tend to be written with their host, without a break separating them. Another interesting regularity she uncovers is the following. If *two* clitics

precede the host, the two clitics tend to be written together, and the host tends to be spelled separately. In order to explain the first pattern, Battisti rightly observes that this tendency can be explained with the crosslinguistic tendency that word-internal vowel-initial syllables are disfavored. The second pattern can also be explained with general principles. Battisti suggests that this pattern is not unlike what happens in contemporary Serbian. In this language, two clitics form a prosodic word together; if there is only one clitic, however, then it constitutes a prosodic word with its host. She proposes, correctly in my view, that this is a consequence of a minimal size requirement imposed on prosodic words. This is true in Serbian, as well as in 19th century Brazilian Portuguese. Battisti's contribution is an interesting attempt to understand some aspects of the orthographic system used in 19th century Brazilian documents. What is particularly appealing about this study is that she tries to explain the attested regularities in terms of general principles of linguistic theory.

Chapter 6 analyses three changes that have applied in the history of Portuguese. These processes are the shortening of geminate consonants, the elimination of the coda consonant, and finally, the fusion of two consonants in the onset. The authors, Hora and Marques de Lucena, argue that these changes conspire to establish a simplified syllable structure in Portuguese. Together, the first two processes eliminate the coda position, which is a case of syllable optimization because the coda is the marked member of a syllable. The third process reduces the complexity of the onsets, which is also a case of syllable optimization. This indicates, according to the authors, that over time, Portuguese has exhibited a consistent tendency towards simplification of syllable structure. It is a case of 'drift' of the type first put forward by Roman Jakobson (Jakobson, 1929). He has argued that Slavic has also exhibited a consistent tendency towards simplification of syllable structure.

The notion of historical conspiracy, or drift, is surely interesting. There is, however, a severe problem with it. It is not entirely clear whether it should be allotted any formal status as an explanatory principle of (historical) phonology. It might be possible that it is just a ghost image projected by the phonologist on the facts. Take Slavic, for instance. According to Jakobson, Slavic has undergone a series of processes, all directed towards the goal of establishing open syllables. However, if that is the case, Slavic must have changed its mind dramatically at some point in its history. All of a sudden, many unstressed vowels were deleted, creating the massively complex syllables that are so characteristic for most of the modern Slavic languages. Russian, for instance, has a word like *vzgljad* 'glance', deriving from historical *vUzUgUljadU*, where capital U marks the lost vowels. How can the abrupt change in orientation be explained if Slavic has a consistent drift towards open syllables? Isn't this drift rather a projection of Jakobson's mind, selecting only those processes that do establish simplified syllables, while not selecting those processes that establish relatively more complex syllables? The same could also be said about Portuguese. If it is really true that the drift towards open syllables plays an active role in the development of Portuguese, then how can this be reconciled with the fact that there are other processes going in the opposite direction? One clear example of this is the reduction of the vowel in a word like *noite* 'night', in contemporary BP. Deletion of this vowel creates a coda, therefore creating a more complex, or marked syllable. What does this mean with respect to the claim that, in its history, Portuguese has drifted towards

simplified syllables? Has it, all of a sudden, altered its course? In my view, contra-drift phenomena can always easily be found, and it shows that the notion of ‘continuous drift’ has no formal status as an explanatory principle in historical phonology.

Chapter 7 deals with a classical problem of Portuguese, metaphony. One of the problems posed by this process is that its application domain is quite restricted. Not all words undergo it, even if the phonological environment is met. The process is also morphologically restricted. For instance, words that do undergo the process only do so in the plural, and in the feminine, singular. An example illustrating the former is *p[o]rco* ‘pig, sg.’ vs. *p[ɔ]rcos* ‘pig, plur.’ The author of this chapter, Moresco Miranda, proposes an analysis using all the power Optimality Theory has developed. The central idea of the analysis is that the trigger of feature agreement, a constraint disallowing certain features to be each other’s neighbors, is relativized to the word edge. This accounts for the fact that the combination [ɔ] + [u] (spelled as <o>) is not allowed in the case of <porco>, so that underlying /pɔrku/ is changed into [porku], whereas underlying /pɔrkus/, the plural, is not changed, and is realized as [pɔrkus]. Of course, this analysis works, but, surely, it does not have very much explanatory depth. Why should it be the case that the word edge is so special, so that the constraint against the two adjacent features no longer holds in that environment? In my view, the topic is very interesting and intriguing, but I have some doubts with respect to the analysis proposed in this chapter.

In Chapter 8, Schwindt develops an analysis of the way two processes, vowel harmony and vowel lowering, interact with deletion of the theme vowel in the verbal system. The intriguing aspect of this interaction lies in its *opaqueness*. That means that the trigger of the two processes is not present at the surface, due to the fact that it has been deleted. This kind of interaction is known as ‘counterbleeding’. An example is the underlying form /segi+o/ ‘I follow’. At the surface, this form is realized as [sigu] (where the raising of /o/ to [u] is caused by reduction of /o/ in unstressed syllables). Importantly, the change to [i] of the first vowel is caused by the theme vowel /i/. Yet, this vowel has disappeared. How can this be possible? Interactions of this type, where the trigger has been removed, are notoriously problematic for Optimality Theory, because this theory is non-derivational, at least in its classical form. OT has therefore been modified in various ways to explain opacity effects. One such modified version is OT-CC, that is, Optimality Theory with Candidate Chains. It is within this theory that Schwindt develops his analysis of the interaction between vowel harmony and vowel lowering, on the one hand, and deletion of the theme vowel, on the other.

Schwindt’s analysis is descriptively adequate. OT-CC does the job! In my opinion, the article has only one shortcoming. In the literature, one alternative analysis has been developed; one that also explains the opacity effect. This is Wetzels’ account published in (1995). Wetzels proposes that the root node of the theme vowel is deleted, but that its aperture features are maintained. These features then dock onto the preceding vowel, raising them. This analysis can straightforwardly be incorporated into standard OT, by applying the standard faithfulness constraints at the level of the aperture node. This approach would explain opacity, without the adoption of candidate chains. Schwindt is aware of Wetzels’ analysis, yet he does not seriously consider it. The following important question, then, still remains to be answered. Do we still need OT-CC to account for the facts of BP if we allow faithfulness constraints to apply at the level of the aperture node?

One of the differences between European Portuguese and Brazilian Portuguese is that the latter vocalizes a lateral in the coda position. In other words, *l* is changed to *w* in the coda position. In the south of Brazil, however, this change has not yet been completed. In this region, it is still an ongoing change. This ongoing change is the subject of Collischonn's study in Chapter 9. Following many other researchers of BP-phonology, she claims that in BP, word final consonants tend to be syllabified in the onset of the next word, if that word starts with a vowel. This resyllabification explains why vowel-initial words tend to block vocalization of the *l* preceding it. It follows from the fact that *l-to-w* vocalization applies only in the coda. This automatically entails that it cannot apply if *l* is located in the onset, due to resyllabification. In her study, Collischon discovers that the nature of prosodic constituents plays a crucial role with respect to resyllabification. Prosodic constituents that are high in the prosodic hierarchy tend to inhibit it, whereas the constituents located lower in the hierarchy tend to allow it. The prosodic phrase, for instance, typically allows it. As a consequence, words that are not separated from each other by an intonational phrase boundary or an utterance boundary do undergo resyllabification. Therefore, they typically undergo vocalization of *l-to-w*. Collischon's result is interesting, because it differs from the results obtained by other researchers. Future research should give more insight into the exact conditions under which resyllabification is allowed at the edges of higher constituents.

In Chapter 10 Cristófar-Silva and Almeida compare the epenthetic vowels of BP with the high front vowels that are supposedly present already at the underlying level. Their study has two important results. First, both types are not always pronounced. Whether they are pronounced is largely determined by syllable structure; both types tend to be pronounced if that optimizes syllable structure. That is, if the result of pronunciation is a CV syllable. The second, very important result is that non-epenthesized high front vowels systematically have a longer duration than epenthesized high front vowels. On the grounds of the first result, the two researchers claim that both types of high vowel are present in lexical representations. In order to distinguish them, fine phonetic detail has to be added to these lexical representations. In their view, this confirms one of the major tenets of Exemplar Phonology, according to which fine grained phonetic information constitutes an essential ingredient of lexical representations.

It might be true that epenthetic vowels are part of lexical representations, in particular if it can be shown that epenthesis in BP is a lexical process. However, I am not convinced that the lexical presence of epenthetic vowels forces us to assume that fine grained phonetic detail must also be part of lexical representations, as Exemplar Phonology would have it. Theoretical phonologists have established some time ago that there is a variety of truly phonological processes distinguishing epenthetic vowels from vowels *of the same quality* that are underlyingly present. A clear example of such a process is stress assignment. Thus, in Alderete (1999), it is convincingly shown that, in many languages, stress skips epenthetic vowels, but not underlying vowels of the same quality. This means that in some way, there must be a *phonological* difference between the two vowel types, at least in some languages. One way to express the difference could be along the lines of Halle and Vergnaud's 1990 essay on stress. They argue that epenthetic vowels are not represented on the grid, whereas underlying vowels (of the same quality) are. Since epenthetic vowels are absent on the grid, they are skipped by the stress rule. The same

strategy could also explain the BP facts. The epenthesized high vowel is not represented on the grid, whereas the non-epenthesized vowel is. This representational difference is fed into the phonetic component, where the two types are allotted different durations, such that epenthetic vowels are assigned less duration than non-epenthetic vowels, all things being equal. The point is that a phonological difference must be made between the two types of vowel. This being the case, it is no longer necessary to incorporate fine phonetic detail into the grammar. Summarizing, the results of Chapter 10 are highly significant, but the radical interpretation proposed by its authors is unwarranted.

In Chapter 11 the author, Bonilha, investigates whether Uffmann's theory of Optimal Geometries (Uffmann, 2003) can account for the patterns found in the acquisition of consonants and vowels by Brazilian children. If I understand Bonilha's argument correctly, she recognizes that the theory does explain the patterns, but nonetheless she criticizes Uffmann's approach because the theory seems to be inconsistent. This has to do with the role GEN plays in Uffmann's theory. She then goes on to propose her own theory to account for the acquisition patterns; a theory which she calls Implicational Demotion of Features.

I cannot agree with this argument at all. Uffmann's 2003 theory of the segment claims that hierarchical structure is very important at the segmental level, a point OT had lost sight of at the time when Uffmann came up with his proposal. Once it is the case that hierarchical structure is crucial at the level of the segment, it is only natural to assign to GEN (the generator) the task of properly organizing a segment's structure. If a segment satisfies the canonical, universal geometry, the candidate containing it is admitted to EVAL (the evaluator). If this is not the case, that is if a segment is not properly organized according to the criteria of the universal theory, then it is not admitted to EVAL. I fail to see why this could be a problem.

In Chapter 12, Matzenauer introduces the concept of *Harmony Scale of Constraints*. With this notion, asymmetries in the acquisition of segment inventories can be explained. This notion is inspired by the Harmony Scales of classical OT. The scales of classical OT are converted into a set of universally ranked constraints. For instance, according to the Harmony Scale governing place nodes, it is the case that the coronal place node is more harmonic than any other place node. From this, a set of markedness constraints is derived. One constraint penalizes coronal place nodes and the other markedness constraints penalize any other place node. Given the Harmony Scale governing place nodes, the constraint penalizing the coronal place node will universally be ranked *below* the constraints penalizing any other place node. The author insists that the type of scale she introduces are not like that. They are present in CON, the component containing the list of constraints.

It is not clear at all to me whether there is a real and substantial difference between the Harmony Scales of classical OT and the *Harmony Scale of Constraints* proposed by the author. Take, for instance, the fact that in BP, stops are acquired before fricatives. To explain this, the author proposes her version of Harmony Scales. Basically she says that not having a fricative is more harmonic than not having a stop. Of course, from this, it automatically follows that stops are acquired before fricatives. As far as I can see, exactly the same strategy can be pursued along the lines of the classical Harmony Scales. One such

scale would say that stops are more harmonic than fricatives. From this, two constraints would automatically be derived. One constraint penalizes stops and the other penalizes fricatives. Given the Harmony Scale, the former is universally ranked below the latter. In my view, this strategy also explains the Brazilian acquisition pattern.

The authors of Chapter 13, Leite and Franchetto, present an overview of the history of the research done on the languages of the indigenous peoples of Brazil. The chapter starts with a brief sketch of the history of these peoples. It is clear that their fate has been quite tragic. Thus, their number declined from about 5 million at the time of discovery by the Europeans to approximately half a million today. Fortunately, however, the interest in the languages of the indigenous people has been steadily growing. Hopefully, this is the first sign that the future of the Indian people in Brazil will be bright. It is very troublesome, though, that many languages of Brazil are on the verge of becoming extinct. According to some pessimists, so the authors explain, eventually all native languages of Brazil will share this fate. They point out that maximal effort should be made to document these languages, not only in the traditional way, but also in a digitalized version, so that documents are maximally accessible.

In Chapter 14, Storto and Demolin test Ramus et al. (1999) hypothesis with data from Karitiana. According to this hypothesis, the rhythm class of a language (that is, whether it is stressed timed, syllable timed or mora timed) is determined only by the proportion of vocalic intervals and the variability of consonantal intervals. Basically, this means that syllable structure is all important, because the proportion of vocalic intervals and the variability of consonantal intervals depends on the inventory of syllable types a language allows. Storto and Demolin show that this prediction is not completely borne out. Karitiana has the same syllable types as Japanese, a mora-timed language. Yet, Karitiana does not entirely behave like Japanese. The authors observe that Karitiana is a mixed language. On the scale from stressed timed to syllable timed to mora timed, it is located in between the last two classes, somewhere in between a language like French, which is essentially syllable timed, and Japanese. They try to explain this with some important phonological properties of Karitiana. Thus, they note that this language has prenasalized consonants, segments that are relatively long, phonetically. It seems to be the case, then, according to these authors, that additional factors have to be taken into account to explain a language's rhythm class.

In Chapter 15, finally, Pisanço investigates creaky voice in Mundurukú, both from a synchronic and a diachronic point of view. He proposes that, phonetically, creaky phonation in Mundurukú is primarily controlled by the thyroarytenoid muscles. On this assumption, the effects creaky voice has on tone can be explained. It explains, for instance, why the language does not allow creaky vowels with high tones. This would directly follow from the interaction between the thyroarytenoid and the cricothyroid muscles. Historically, creaky voice is a consequence of the loss of a glottal stop in certain positions. After the loss of the glottal stop, the glottal constriction shifted to the neighboring vowel, resulting in creaky voice. Interestingly, in modern Mundurukú, there still is a contrast between vowels with creaky voice and intervocalic glottal stops. This contrast manifests itself phonetically as a timing difference. In vowels with creaky voice, glottal activity is in the vowel itself, whereas in the case of glottal stops, it is mainly located at the transition from one vowel to the other.

It should be mentioned that this article is very difficult to read because the phonetic symbols have been severely distorted. As a result, crucial illustrations are sometimes impossible to evaluate. Clearly more attention should have been paid to the final layout of this paper.

References

- Alderete, J. 1999. Head dependence in stress-epenthesis interaction. In B. Hermans and M. van Oostendorp (eds.) *The derivational residue in Optimality Theory*. Amsterdam: Benjamins: 29–50.
- Halle, M. And J.-R. Vergnaud 1990. *An essay on stress*. Cambridge, Massachusetts: MIT Press.
- Hayes, B. 1995. *Metrical stress theory: principles and case studies*. Chicago and London: University of Chicago Press.
- Hyde, B. 2001. Metrical and prosodic structure in optimality theory. PhD Dissertation, Rutgers University, New Brunswick, New Jersey.
- Jakobson, R. 1929. Remarques sur l'évolution phonologique du russe comparée à celle des autres langues slaves. In *Travaux du Cercle Linguistique de Prague*. Reprinted in *Selected Writings*, Volume 1.
- Ramus, F., M. Nespore and J. Mehler 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265–292.
- Uffmann, C. 2003. Optimal Geometries. Talk presented at OCP1, University of Leiden.
- Wetzels, L. 1995. Mid-vowel neutralization in the Brazilian Portuguese verb. *Phonology* 12: 281–304.

C. W. McClelland (2008) ***A Phonology of Tarifit Berber***

Lincom Europa: Munich. Series: LINCOM Studies in Afroasiatic Linguistics (xxii + 154 pp., Paperback, ISBN 978 3 89586 129 1, Price: 66.20 Euro/USD 92.68)

Reviewed by:

Judith Rosenhouse

Swantech – Sound Waves Analysis and Technologies Ltd. Haifa, Israel
swantech@013.net

Dialects of the Berber language spread out in Morocco, Algeria and a little in Libya, to the sub-Saharan countries of Mali and Niger. This language has been mainly oral, so that few texts have been available for study. Tarifit is the north-eastern dialect group of the three large Berber dialects in Morocco. (The other two are Tamazight and Tashelhit). In Northern Morocco also Spanish and Moroccan Arabic exist and interact with the Berber dialects. The Moroccan government has been encouraging vernacular literacy, involving the representation of Berber in some written form. To this end, a phonemic study of the language is necessary and this book takes one step in this direction.

C.W. McClelland III has been studying the Berber language Tarifit for a considerable period. During this time, he has published several books on this language, including a dictionary (2004), texts (2008) and a study of the relationships between syntax, narrative structure, and prosody (2000). McClelland's present book focuses on the phonology of this little-studied language from primarily a functionalist approach.

The main audience is that of linguists who are familiar with phonological principles and practices and their students. Therefore, the text is explained succinctly and refers the novice reader to footnotes which explain linguistic terminology and notions.

The book is made up as follows: 1. Introduction, which presents a general background of the language, with two maps of the spread of Berber in Morocco and in North Africa, and includes a description of its morphology (verb structures and nominal structures and their affixes). The main part of the book: 2. The Phonology. In this part, the methodology is first explained and is followed (section 2.2) by a discussion of the main distinctions between units and sequences in this language. Three groups of definition problems are described: that of [dʒ], that of labialized consonants (e.g. [m^w, g^w]) and that of consonant gemination. Also, the question of vowels vs. consonants (in u/w, i/j) is discussed. Section 2.3 describes the underlying units and allophonic variations of the consonants and vowels in this language also referring to the status of schwa (which is placed so that it will break up consonant clusters but not between the consonants of a geminate). Section 2.4 discusses the gaps in the segmental analysis, which are mainly due to the linguistic phonemic or phonetic description of the units. For example, it seems that pharyngealization, labialization and gemination work together to form the present tense in certain verbs, but the stems of such verbs appear to have already the labial phoneme /w/. In other cases, phonetic features of some phonemes include both pharyngealization and labialization, in the same environment as only pharyngealization (example 200, p. 105: /ʃ^ɣʃ^war^ɣ/ 'be filled' vs. /ʃ^ɣʃ^war^ɣ/ 'advice'¹). Section 2.5 (which is missing from the list of contents) contains four tables of the Tarifit consonantal and vocalic phonemes, including a long table which summarizes the distinctive features of the consonants. Section 2.6 moves to examine syllable structures. There are several structures which are "non-suspect", such that they have a vowel as the peak (beginning with V, and going up to CVCC), but there are some other patterns (CC, CCC) where schwa is inserted to enable articulation. Section 2.7 ends this chapter with consideration of the phonological word, shown by three tables of word templates arranged by frequency of occurrence in his data. Some of the conclusions are that template II allows no open syllables, whereas Template I and III never had a schwa in an open syllable; a geminate must be preceded by a vowel (either schwa or not); if there are three consonants in sequence, then either a schwa was omitted in the transcription or one consonant is syllabic; no vowel sequences are allowed; there are two basic phonological words: (a) those with alternating sequences of CV and CVC patterns, and those of CVC sequence. Both patterns allow a consonant cluster at the end of a word (p. 118).

The book ends with a 2-page Conclusion (pp. 119–120) which notes that the study has found 38 phonemes: 4 labials, 19 alveolars and alveo-palatals, 12 velars, uvulars,

¹ McClelland comments that these may be merely homophones. However, since the second word, 'advice' is borrowed from Arabic, this difference apparently distinguishes it from the original Berber word [JR].

pharyngeals and glottals, and three vowels (/i, a, u/)² (Table 21). The author describes the “imbalances” in the list, such as no bilabial unvoiced phoneme parallel to /β/, though /f/ apparently serves this role, and there is no velar equivalent to /n/ though there is a labialized velar nasal phoneme /ŋ^w/. The author is puzzled also by the existence of labialization in only eight consonants and suggests the presence of dialect mixtures incorporating this phonetic feature. More research is needed, writes McClelland, on this issue, as well as to study labialized + pharyngealized phonemes, such as /ʒ^{fw}/, /ʃ^{fw}/).

An Appendix at the end of the book lists the 551 word corpus on which the syllable and word analysis is based (pp. 123–146), followed by a selected bibliography (pp. 147–149) and an index (pp. 151–154).

In sum, this small book contains much to ponder for experienced linguists and phoneticians, as well as for students. The questions raised by the author following the analysis are mainly of a phonetic and not phonological nature, and are therefore salient for us, but the answers may involve features borrowed from other Afroasiatic regional languages.

References

- McClelland, III, C.W. (2000) *The Interrelations between Syntax, Narrative Structure and Prosody in a Berber Language*, New York: Edwin Mellen Press
McClelland, III, C.W. (2004) *A Tarifit-English Dictionary: Documenting an Endangered Language*, New York: Edwin Mellen Press.
McClelland, III, C.W. (2008) *Texts for Analysis*. Munich: Lincom Academic Publishers

Revue Parole, Issues 41–42 and 43–44 (2007)

The University of Mons-Hainaut, Belgium, editor: Marielle Bruyninckx (150 pp. including Editorial Introduction, articles, book reviews; ISSN 1373-1055, price per issue: in Belgium 13.59 Euro, elsewhere: 16.07 Euro; per year: in Belgium 45.7 Euro, elsewhere: 51.2 Euro)

Reviewed by:

Judith Rosenhouse

Swantech – Sound Waves Analysis and Technologies Haifa 32684, Israel
swantech@013.net

“Revue Parole” is not the only linguistic publication published within the walls of a university, but the two brochures which arrived at our desk for 2007 include some interesting papers for our readers. Moreover, the editor, professor Bruyninckx, informs the reader in her editorial that the journal has reached its 10th anniversary. On this occasion, an index of all the publications in this period was published as well. We review the papers in the brochures from 2007 in the order they are published. In issue “2007–41/42” we find three articles from different fields:

² In addition, there is the central schwa, which is non-phonemic and is not mentioned in this listing.

Nadia Duchêne, “Langue métissée et traduction” (pp. 3–23)

This paper analyses the translation of the French novel “Texaco” by Patrick Chamoiseau (appeared in 1992 and received the Goncourt Prize) into Spanish. The novel reflects the history and development of Martinique and is written in a Creole style, which differs from regular French, and is therefore modified in the original French book into a somewhat “creolized” French version by various syntactic, morphological and phonological features. This increases the difficulty for the Spanish translator (E. Calataud), for she had to transfer not only the contents into the target language, but also the cultural and linguistic features of the origin. Among the different examples, Duchêne lists some “marks of orality” in slang expressions, which are written and faithfully translated (e.g., French “floup” – Spanish “flup”). The author concludes this paper by calling for further study of this kind of literature and its translation being a fertile field of linguistic observations.

Pauline Sirois and Andrée Boisclair, “Les débuts de l’apprentissage de la lecture chez l’enfant sourd : présentation d’un modèle d’intervention pédagogique” (pp. 25–77)

The fact that hearing impaired children are at a phonological disadvantage compared to hearing children is obvious when they begin to learn how to read and write at school. Therefore, hearing impaired children need to have special learning programs. This paper suggests a new methodology for teaching these children how to read, putting stress on the search for meaning. At first, the authors suggest that the child should understand the link between writing and reading, by using pre-school writing (or scribbling), which is a cognitive process of connecting signs with sounds and meanings (logographics), and developing it into the systematic phonological units of speech and writing. This initial stage is followed by the “alphabetic” stage and then by the “orthographic” stage (Firth, 1985). These three stages are expressed as six different lessons. Another theory they follow is by Ehri (1999), which adds a “semi-phonetic” (or partially alphabetic) stage between Firth’s first and second stages. The authors’ pedagogical model combines these theories to suggest three steps: 1. Conceptualization of the alphabetic principle in invented writing and development of a first visual vocabulary with graphic-phonetic indices. 2. Development of assembling procedures (phonological mediation). 3. Reduced dependence on the use of assembling procedures (emergence of orthographic vocabulary). The authors add that this model has been experimentally used in recent years with hearing impaired children in pre-school and elementary school grades. The results of these experiments showed that the learning route and results of these children were similar to those of age-equivalent hearing children. Continued research is needed to prove that advantages of this method for hearing impaired children.

Monique Vion and Annie Colas, “Signaler « à la volée » les ruptures et les frontières prosodiques : une technique comportementale” (pp. 79–119)

The paper presents an on-line paradigm designed by members of the lab team in the University of Provence to study the perception of breaks and discourse boundaries by naive and expert listeners. The authors used four versions of a 1 minute long spontaneous speech sample: original, low pass filtered, re-synthesized, and inverted (phonemes exchanged with others by certain rules). These variations enable testing various degrees of perceptual difficulty. Sixty naive listeners (15 listeners x 4 versions) listened in a sound-proof room to a single version of the short text and had to press a key when they heard a

break (or limit) between words. This part required quick on-line responses. Then the original text was administered to two expert listeners for an off-line analysis (focusing on types of intonation contours and the degree of separation between words). The paper presents many tables, graphs, and spectrographs and an appendix presents the studied text versions (utterances per versions). The results of these psycholinguistic tests show relatively good agreement between the listeners and the experts (40% similarity in the major intonation breaks). Breaks were marked mainly at intonation boundaries. The naive listeners marked the breaks before the following silent pause when the signal offered a rapid and wide rising pitch in the higher part of the speaker's range, or a steady pitch in the lower part of his speech range. The results also suggest that in this tool, which requires rapid speech perception, the listener's main resources are directed to grammatical and discursual goals rather than to prosody, but it also enables bringing the integration process of prosodic information closer to speech perception. The results also show significant differences between various intonation contours (rises or stationary) and the following silence breaks and appear to constitute attention cues (also apparent in maternal language).

Issue 43–44 (2007) has the following four papers:

Muriel Barabzan, “De la grammaire explicite à l'intériorisation : une question d'adéquation linguistique, didactique et cognitive” (pp. 151–182)

The question dealt with in this paper relates to the manner of learning an L2 (in this case French as the L2). The main point of view is that the learner has to internalize the information he learns. This really means using metalinguistic skills (and gestalt notions) which the learner has previously acquired (with the L1) and which are used in explaining the learned material. The method of teaching the past tense in the verb system in several textbooks is then examined from the perspective of mental processes leading to the development of the learner's interlanguage. This paper, though interesting in itself, does not really touch upon phonetics or phonology.

Martin Howard, “Sociolinguistic variation in contemporary French: Insights from the spoken language of the media” (pp. 185–214)

The aim of this paper is to examine variations in the occurrence of the liaison feature as a sociolinguistic characteristic in French, including differences between the French used in France and that used in Canada. This topic is studied using the language of the media language (TV and radio programs) for language samples. Liaison has been reconfirmed as a feature of high prestige enthusiastically adapted by journalists irrespective of their sex and reporting role, and is used much more than in a native speakers' daily speech. This is related to the higher degree of formality in media language, as well as to some “trendiness” attached to liaison. Liaison also seems to require a higher literacy level, which may be one of the reasons for its limitation to certain morpho-syntactic contexts. However, the author has found also hyper-correction errors in its use. In addition, liaison has been found more in male journalists' formal speech than in female reporters' speech. This relation of liaison occurrence is also present in the less formal style of men's speech. The author suggests that further research is needed into the differential interaction between style and gender in everyday spoken French. In general, French liaison has been decreasing since about 1960,

as mentioned in the literature, and this is also noted in this study. The author finally calls for further research of other sociolinguistic features of liaison in other media programs, including less formal and more local ones, considering that media language is an important discourse type for the study of sociolinguistic variation.

Laurent Lefebvre, “Etude des aptitudes langagières chez les patients atteints de la maladie d’Alzheimer” (pp. 217–238)

This author investigates 30 Alzheimer patients in the three stages of the disease (1st, 2nd and 3rd). These stages mark the beginning, the development and the most severe stage of this disease as described in the literature. However, in contrast with the literature which usually focuses on behavioral aspects of such patients, Lefebvre examined the phonological and phonetic difficulties as they changed in each stage. The tests he made were based on the aphasia examinations of Ducarne de Ribaucourt (1989). The results are presented in terms of grading (per 100) and clearly show the gradual deterioration of expressive skills, oral comprehension skills and reading and writing skills in these patients. Tasks on the tests included repetition of syllables, words and phrases, as well as the production of speech describing pictures, writing spontaneous and dictated texts, etc. The author suggests that this method of linguistic testing may contribute to better understanding of the nature of the cognitive and linguistic problems of Alzheimer patients.

Fanny Meunier, Michel Hoën, Juan Seguí, “L’effet de la fréquence des mots en modalité auditive : approche expérimentale” (pp. 241–262).

Word frequency has been studied visually (in written texts), but the authors of this paper examined word frequency in the spoken mode. Literature on this topic has usually examined monosyllabic words. This paper reports experiments with long and morphologically complex words, i.e., those that include prefixes and suffixes. The participants had to complete decision tasks concerning the roots with different frequencies and the frequency of words with the different affixes. A clear frequency effect was found for these words, as listed in the appendices. The authors suggest that this shows the role of surface frequency in the auditory mode during word identification and recognition.

To sum up: The topics of the papers are varied and deal with up-to-date topics which are related to phonology or phonetics. The papers in these two brochures primarily sum up the results of their various tests, partly obtained by statistical analysis, although not many examples of the material are given. Each paper ends with abstracts in both French and English, which helps readers for whom French is not the mother tongue. In addition, the editor notes in her Introduction that the external format has been changed in this volume of 2007. This “university journal” seems, then, to be gradually maturing, on the way to establishing itself as a journal for researchers from all over the world.

References

- Chamoiseau, P. (1992) *Texaco*, Paris: Gallimard
Chamoiseau, P. (1994) *Texaco*, Barcelona: Anagrama. (Translator: Emma Calatayud)
Ducarne de Ribaucourt, B. (1989) *Test pour l’examen de l’aphasie*, Paris: Editions et Applications Psychologiques

Jörg Peters (2006)

Intonation deutscher Regionalsprachen

Berlin/New York: Walter de Gruyter Verlag (Linguistik – Impulse und Tendenzen; 21, XIV + 510 pp. including Preface, Abbreviation Index, Bibliography, Author Index, Appendix, ISBN 978-3-11-019096-0 Price: Euro 108, \$ 173)

Reviewed by:

Verena Teschke

Fachbereich II Germanistik – Ältere deutsche Philologie – Universität Trier, Germany
teschke@uni-trier.de

Even though it is widely believed that the prosodic characteristics of a dialect belong to its most striking features, intonational variation has received rather little attention in classical German dialectology. But the present-day situation offers far wider possibilities than ever before to close this gap. So the amount of research in this field has remarkably increased during the recent years.

The book “Intonation deutscher Regionalsprachen” by Jörg Peters does not provide, as the title might suggest, an overview of the intonation of German regional varieties. But after all, that’s not what it is intended to be. The author’s main concern is the development of a system to describe and analyze Standard German intonation that is easy to understand and can also be applied to regional varieties. Furthermore, he seeks to build up a theoretical framework that can be used to compare the intonation systems of these varieties.

The book itself is divided into two parts, preceded by a short introduction. Part A (Chap. 2–4) provides the theoretical background. Chapter 2 starts with a brief historical review and summarizes some of the previous research concerning intonational variation in the German language area. As his main point of criticism about these early studies, the author names a lack of ‘structural information’ and claims that such information cannot be extracted from describing individual intonation contours, but only from studying intonational variation systematically. This leads directly to today’s ‘state of the art’ and to Chapter 3, which is concerned with the phonetics and phonology of intonation. This chapter gives an overview of general ideas and the historical development of autosegmental-metrical phonology and offers a short introduction to intonational analysis within this framework.

Chapter 4 sets the focus of the study on the identification of systemic differences and reflects on possibilities to compare intonational systems. This section is very important, because the author specifies his own assumptions and raises the questions to be answered in the book. As a starting point, he asks under which conditions two intonation contours or intonation systems can be considered as different, from a phonological point of view. At first sight, one may think it would be sufficient that they have different inventories of tonal units, but then there follows the question whether those differences (e.g., in timing) are really systemic and therefore phonological in nature or whether they are simply phonetic. The author then reflects on the notion of phonological equivalence and argues that for the comparison of paradigmatic tonal contrasts, it is crucial that the tonal units in question not

only share the same distinctive features, but also occur in the same structural position within a given contour.

He illustrates this point by presenting some fictitious examples. If the comparison is simply based on the comparison of ToBI-like (Tones and Break Indices, see Beckman & Ayers-Elam 1997) annotation systems, a problem rises since equal notation does not guarantee phonological equivalence. On the other hand, different notations do not automatically mean that there exist systemic differences. It is essential that the distinctive (phonological) features that constitute the tonal contrast are the same in both varieties and that the opposition occurs in the same structural position within an intonation phrase (IP). Consequently, the author claims that a paradigmatic comparison of tonal contrasts is only possible with respect to the complete intonational systems.

The second and much longer Part B encompasses chapters 5–12 and presents the application of the general ideas developed in the previous chapters. Chapter 5 describes the intonation of Northern Standard German (NSD), based on earlier works, by Pheby (1980), Uhmann (1991), Féry (1993), Grice & Baumann (2002; GToBI), but with a few modifications. This description serves as a reference for the analyses of the regional varieties and introduces all the necessary conceptual categories. Chapters 6–11 contain the analyses of the intonational systems of six German regional varieties, the urban vernaculars of Hamburg, Berlin, Duisburg, Köln, Mannheim and Freiburg. The speech data were derived from six different corpora, all of which contain natural speech, mostly from informal conversations between ‘native speakers’. The audio files (*wav) of all parts of speech that are described and analyzed are provided on the accompanying CD and many figures, schematic representations of contours or F_0 plots and waveforms drawn by the PRAAT program are also included, especially in the analytic chapters.

These chapters are all equally structured. They start with an introduction to the characteristics of the respective variety. The major part describes the basic inventory of tonal units, consisting of nuclear and prenuclear pitch accents, IP-initial and IP-final boundary tones, as well as their possible combinations. A table finally lists all contours that were found in the speech data as schematic representations. Also added are so called ‘Finite State Grammars’, from which every possible contour can be generated by means of specific phonetic implementation rules. Based on this, it is possible to identify all tonal contrasts that are realized with respect to structural position. But the author points out that in view of a comparison of intonation systems, it is not sufficient that they have the same tonal contrasts and are therefore phonologically equivalent. The varieties should also make use of these contrasts in the same way; they should be semantically equivalent, meaning that they have the same conversational function. So another step, the analysis of ‘semantic features’, follows. These features are carried by assumed ‘tonal morphemes’ and are very abstract. The contours are classified according to their typical use in conversation. Concerning the whole IP, the choice of boundary tone (or its absence) defines the ‘conversational completeness’ of the utterance. Regarding the accented unit, the presence or absence of a trailing tone involves the ‘informational completeness’ of the accented part, while the choice of accent tone defines its ‘communicative relevance’. For the prenuclear accents, the presence or absence of a trailing tone involves the ‘informational completeness’ of the accented part, while the choice of accent tone indicates the ‘availability of information’.

After the treatment of the ‘simple contours’ follows a part entitled ‘extended grammar’ that deals with so called ‘modified contours’. It has a closer look at the use and function of some phenomena that have frequently been investigated in intonation research, such as downstep, contrastive accent, late peak, ‘calling contours’ and clitic IP’s. At the end of each chapter, speaker-specific characteristics are presented in the section which explores differences between speakers with a ‘close to standard’ segmental level and those who clearly speak a dialectal variety. Some biographical information on the speakers, as well as information about their segmental dialectal level, which served as the basis for group assignment, can be found in the appendix.

Chapter 12 finally presents the synopsis of the preceding analyses. The intonational systems of the six varieties are compared to each other as well as to the Northern Standard German (NSD) system. First, the schematic illustrations of the nuclear contours are put together in a table, presenting those that correspond to the contours of NSD with respect to tonal contrasts and semantic function in rows. In order to further ensure the functional correspondence, the frequency of use has also been investigated. At first sight, it seems that the differences between the studied dialects and the NSD system increase from north to south. The Hamburg system does not show differences from NSD, but the Köln, Mannheim and Freiburg systems look almost completely different from it. But following the argument that tonal contrasts are to be compared only with respect to the complete systems, the ‘Finite State Grammars’ of all varieties are then taken into account. This shows that some of the seemingly systemic differences are simply phonotactic in nature. In other cases, contours that are usually annotated identically do not have the same phonological status within their respective systems. Usually, a high tone is represented by “H” and a low tone by “L”. For example, the nuclear falling contour of the NSD system is annotated as H*LL%. But this notation contains redundant information, as the height of a trailing tone in the NSD system is always predictable with respect to the accent tone: H* is always followed by a low tone. Therefore, tonal height in the position of a trailing tone is not distinctive. In such cases where there is no tonal contrast realized, “T” (for “tone”) can be used instead, so the nuclear falling contour in NSD is notated as H*TL%. This is also appropriate for all other varieties in question except for the Berlin vernacular. There, in the structural position of the IP-final boundary tone, tonal height is not distinctive, but it is predictable, because it has the same height as its preceding tone. The redundant notation for the Berlin nuclear falling contour is therefore also H*LL%, but in fact it has to be assumed H*TT% instead, which indicates a systemic difference.

In summary, the book offers an interesting new perspective, not so much for beginners, but for readers who are already familiar with the topic. It is therefore especially of interest to researchers looking for an alternative way to describe and analyze intonational systems, as well as to compare them and to deal with intersystemic variation.

References

- Beckman, M. E. & Ayers-Elam, G. M. (1997) *Guidelines for ToBI Labelling* (version 3.0, March 1997). The Ohio State University Research Foundation, unpublished manuscript, online: http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/.
- Féry, C. (1993) *German Intonational Patterns*. Tübingen: Niemeyer (Linguistische Arbeiten; 285).

- Grice, M. & Baumann, S. (2002) Deutsche Intonation und GToBI. In: *Linguistische Berichte*, 191, pp. 267–298.
- Pheby, J. (1980) *Intonation und Grammatik im Deutschen*. 2., durchges. Auflage. Berlin: Akademie-Verlag.
- Uhmann, S. (1991) *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer (Linguistische Arb.; 252).

Sara, S. I. & Mauck, S. M. (Eds. and Trans.)
Yehuda ben David Hayyuj, Tanqiyt': A Treatise on Hebrew Vowels,
Translation, Notes and Analysis.

(München: Lincom Europa, xxi + 154 pp., hardback, ISBN: 978-3-89586-312-7,
 Price: Eur 113,80/£ 81,90/\$ 167,30)

Reviewed by:

Tamar Zewi

Department of Hebrew Language, University of Haifa, Israel, tzewi@univ.haifa.ac.il

The new translation of ‘the book of punctuation/pointing’ is intended, as indicated by the editors and translators, for readers from various disciplines including general linguistics and other related fields, and not necessarily specialists in Arabic and Hebrew. Accordingly, the new edition very conveniently presents the original Arabic text followed by its well-known Hebrew medieval translation by Ibn Ezra on the left side of each page and the new English translation on the right side. As stated on p. xxi, the Arabic text and the Hebrew translation are based on earlier printed editions, those from 1879 and 1844 respectively, and the current edition also gives vocalization marks in the Biblical Hebrew examples cited in the Arabic and Hebrew texts. The transcription of Arabic and Hebrew words is not the one common among Semitists, but is based on international phonetic conventions accepted in general linguistics (a transcription chart, though only for the Arabic letters, is given on p. x, whereas the Hebrew consonants and vowels are eventually explained in charts on pp. 41–44. The reader is then left to work out the transcription of the Hebrew letters and the vowels of both languages.)

This practice serves as long as the transcription is consistent and reasonably follows general linguistic conventions, which generally seems to be the case in this book, but there are some inconsistencies, as indicated below. The choice of the transcription of the geminate and non-geminate בגדכפ"ת, for example, *x* for כ and *k* for כ, or *v* for ב and *b* for ב, does not help demonstrating each pair as two allophones of one phoneme respectively.

The book has been made ‘user-friendly’ and more transparent to potential non-Semitist readers by adding titles to the chapters and by maintaining all the Arabic terms in transcription in their proper place inside the translation, and giving their English translation alongside in inverted commas. Though the reading is less fluent this way, it allows the reader better to follow the Arabic original and to learn the immediate meaning of the text. Each chapter contains comments in a separate section and concludes with a more general analysis of its content in a special part dedicated to analyses.

As for the book's overall layout, a general introduction (pp. xi-xxi) precedes the text and briefly describes the linguistic background for the text and related medieval and modern scholarly studies. This description is sometimes oversimplified, e.g., Ben Saruq's dictionary is described as listing roots from one to five letters (p. xviii). Such an account of the organization of that dictionary is controversial and other interpretations are possible. The text itself is presented on pp. 1–35 and is followed by separate analyses of its 11 chapters (pp. 37–99) and a very short general summary and conclusion (pp. 101–102). Next are an index of Biblical quotations in Hebrew letters and in transcription, and their English translation (pp. 105–113), an index of single Hebrew words mentioned in the text in Hebrew letters, transcription, and an English translation (pp. 115–121), an index of these Hebrew words, sorted by transcription (pp. 122–128) and by English translation (pp. 129–135), a list of Arabic technical terms (pp. 136–139), a list of these terms, sorted by transcription (pp. 140–143) and by English translation (pp. 144–147), a list of Hebrew technical terms (pp. 149), and a bibliography (pp. 151–154).

Since there is an earlier English translation of this text by John W. Nutt (1870), the following discussion of the new English translation presented in the current book is evaluated in comparison with that one. First, it is very useful that the new translation in this book (versus the old translation) cites the Arabic terms in transcription, followed by their English translation in inverted commas. Second, frequently this new translation seems more faithful to the Arabic original than the old one; e.g., it transcribes קמץ קטון (*qamas^f qat^fon*), found in the original in Hebrew letters, as it is, while the earlier translation replaced it with 'tseri' (p. 2). The latter Hebrew term is indeed the common one for this vowel, but it is not the one attested in the original. The same is true for the term פתח קטון (*pataḥ qat^fon*) and its retention in transcription in the new translation (p. 3), while the older translation has the term 'seghol'. Also helpful is the insertion of the quality of these vowels in brackets next to the terms, e.g., קמץ קטון (*qamas^f qat^fon*) is marked [e] and פתח קטון (*pataḥ qat^fon*) [ε]. Regrettably, however, these insertions are not explained anywhere, and the reader, not necessarily a linguist, must comprehend their role with no explanation.

The new translation is clearly far more literally accurate and complete than the former one in many other cases, and it takes no shortcuts. For example, as against an indication of only one Hebrew word followed by 'etc.' in Nutt's translation, the new one gives a full list of the Hebrew words mentioned in the Arabic original, transcribed and translated into English, and followed by the full accurate translation of the discussion found in the Arabic original (e.g., p. 6 and many more examples later on).

All the practices mentioned so far are satisfactory, but the decision in the new translation to render the Hebrew words given as examples in the Arabic text in English seems superfluous, since their meaning is irrelevant to the grammatical text. Moreover, the English translations of the Arabic grammatical terms are not always better than those employed in the earlier translation and are sometimes too literal. For example, terms like *'ism* and *fi'l* (pp. 3 ff.) are literally translated by 'name' and 'action' respectively, although the common word categories 'noun' and 'verb' found in Nutt's translation have long been the accepted terms in the Arabic, Hebrew, and western grammatical traditions. Such renderings could perhaps find their place in a comment explaining the original meaning of the names of the word categories, and not in the translated text itself. However, 'noun' and

‘verb’ are presented as optional on p. 52. Likewise, rendering *fiʿl mādʿiy* by ‘past action’ in the current translation (pp. 5 ff.) is not superior to Nutt’s ‘preterite’, and ‘past tense’ is perhaps more adequate in this case. Furthermore, rendering *lahn* as ‘melody’ in the current translation (pp. 7 ff.) is not better than Nutt’s ‘accent’, both of which could well be substituted by ‘stress’. Similarly odd are the translations of *ʿabwāb* by ‘cases’ (p. 11) and *bāb* by ‘class’ (p. 13) instead of perhaps ‘chapter/s’ or ‘section/s’, especially as the translation ‘chapter’ serves at chapter headings. Strange too is the rendering of *masʿdar / masʿadir* as ‘root/s’ instead of perhaps ‘verbal noun/s’ or ‘primal form/s’ (pp. 11 ff.) and *wazn* by ‘measure’ instead of perhaps ‘pattern’ (pp. 13 ff.). *yuqās* and *qiyās* should probably better be rendered ‘compared/by comparison’ or ‘by analogy’ than ‘patterned’ (p. 29) or ‘pattern’ (p. 31). *ḥarakāt* should probably be rendered by the more established term ‘vowels’ instead of ‘motions’, which is the literal meaning of the word (pp. 40 ff.). The latter ‘vowels’ is indeed displayed as optional on p. 48.

Other translations of Arabic terms are not always consistent, e.g., *wajh* is translated once as ‘facet’ (p. 11) and on other times as ‘surface form’ (pp. 18 ff.). The explanations presented in the analysis of the chapters are not always clearly explaining the issues. Thus, the accentual marking of the end of verses on pp. 71–73, mixes accent marks introduced by the Masoretes with the Modern sign [:]. Another such inaccurate phrasing is the indication that *ḥidʿāfah* ‘annexion’ is a construct state (p. 77), whereas a construct state is the state of the first member of Hebrew annexion, i.e., the *nomen regens*, as correctly stated subsequently on p. 78. References to the bibliography (pp. 151–154) appear only in the introduction to the book (pp. xi–xxi), and are not given in the remarks and discussions in the footnotes and the analyses that follow.

Among the numerous Hebrew examples given in the Hebrew alphabet, many orthographic errors occur – in fact too many to detail here. Some of them involve similar letter forms of the Hebrew letters (mixing between e.g., *h/h*, *d/ḏ*), some reflect similar pronunciation, e.g., מֶרְקָבָה > מַרְקָבָה (*merqava* > *markava*, p. 5); but most often they mix vowels. Especially frequent are errors of the *ḥatʿaf pataḥ* (◌) vowel which is missing or exchanged with *pataḥ* (◌), even in the term *ḥatʿaf pataḥ* itself. Also missing or superfluous are the gemination mark (a dot at the mid line level following the letter), confusion of *pataḥ* /a/ with *qamas* /ɔ/, *tseri* /ɛ/ with *seghol* /e/, *ṣuruq* /u:/ and *qubus* /u/, short with long vowels, etc. It is a pity that such errors occur, because vowel punctuation is the core of the book.

To sum up, this book is interesting reading for linguists, mainly phoneticians, who are interested in historical aspects of Hebrew and Arabic and in the Hebrew phonetic system, and particularly in its vowel marking which links speech with writing. Readers who know these languages can learn much about the background of the system and its development, whereas readers who are not experts in these languages can benefit from the translation without referring to the Hebrew and Arabic parts. The book has not received much research attention until now, compared to other books by the author, *Yehuda ben David Hayyuj*, and the editors-translators are to be recommended for adding this book to the modern library of Hebrew phonetics. However, a corrected edition is really necessary, in order not to confuse learners and readers.

D. Gardner Bonneau and H. E. Blanchard (eds.) (2008)

Human Factors and Voice Interactive Systems

Springer, Signals and Communication Technology Series, Second Edition (469 pp.)

Reviewed by:

Gábor Olaszy

Research Institute for Linguistics, Budapest, Hungary, olaszy@nytud.hu

This book represents one of those publications which focuses on human-computer interaction – a topic that is becoming one of the most important components of our personal life. We must live together with computers and we have to use them (i.e., “communicate” with them). Although developers make efforts to design user-friendly dialogues between the machine and the user, the result is far from being an intelligent, flexible and natural solution. Researching human behavior and reactions to machine-generated messages gradually helps change this situation. This book highlights the importance of human factors in speech technologies, mainly speech-enabled applications. The previous edition with the same title was issued in 1999. The publication of this second edition shows the success of the former volume. The editors themselves concluded that for this book, another collection of research reports are needed, following the changes in speech industry since 1999. Editors organized the content of the book into two general sections. The first nine chapters (from a total of 15) are devoted mainly to interactive voice response (IVR) technology, including the problem of multimodal interfaces. The second part covers a variety of topics, focusing mainly on human behavior and its research from the point of view of speech technology.

The first chapter, by Bernhard Shum, shows the problematic points in IVR usability engineering. He begins by describing the principles and guidelines, showing the limitations of speech recognition, spoken language and cognition. Based on these findings, he shows the best practices for IVR design and gives a short list of guidelines. In the second part of this chapter, he writes about data-driven IVR usability engineering based on end-to-end calls. Finally, an evaluation of IVR cost-effectiveness is discussed.

In Chapter 2, Susan Boyce updates her studies of the natural language system from the first edition with a new title, “User Interface Design for Natural Language (NL) Systems: From Research to Reality”. As she puts it: “There has been quite a lot of industry attention paid to the use of natural language technology in call center environments” since 1999. This technology evolved from research-based lab prototypes to large-scale call center deployments. The NL interface gives freedom to the user to use not only a restricted set of command words, but complete sentences. In other words, the system is prepared to recognize many variations of callers asking for the same service and route them to the appropriate next step in the dialog. The author gives a full overview of NL design, pros and cons, steps of development, user expectations, repair of recognition errors, use of re-prompts for repairing recognition errors, and the design of turn-taking.

Osamujimen Stewart and Harry Blanchard continue the focus on telephone-based IVR systems in Chapter 3 under the title: “Linguistics and Psycholinguistics in IVR Design”. The goal is to show how these two disciplines can help optimize the design of IVR systems for usability and performance. They introduce some new dimensions in the interface

design, such as call flows, the importance of phrase structure diagrams, natural language understanding (grammar), and the relevance of lexical semantics for labeling utterances in natural language categorization techniques. They also discuss the concept of “structural simplification” as it relates to the ways in which humans may use different speech strategies to interact with machines in contrast to interactions with other humans.

The title of Chapter 4 is “Designing the Voice User Interface for Automated Directory Assistance”. Amir Mané and Esther Levin focus on the challenges that one encounters when building a voice user interface (VUI) for automated directory assistance. This is a difficult task, considering the nature of names, addresses, place names, the amount of data, database organization and the fact that telephone directories are designed for humans and not for machines. Several points are mentioned in this respect, like abbreviations, typographical errors, design for optimal optical search, variability in name and acronym pronunciation, and the processing of business names (Xerox, FedEx). At the end of the chapter, general dialogue design issues are summarized.

The next two chapters concern interactive speech systems for mobile and other small consumer devices. In Chapter 5, Dragos Burileanu introduces current trends in embedded systems for these devices. The title is “Spoken Language Interfaces for Embedded Applications”, and the author first evaluates the benefits of spoken language interfaces, and then analyzes the current theoretical and practical solutions. The development of an optimized embedded text-to-speech synthesis system is also described.

Chapter 6, “Speech Generation in Mobile Phones”, is presented by Németh, Kiss, Zainkó, Olaszy and Tóth. They introduce general concepts and application areas involving speaking through mobile phones. The small-scale computing device in mobile phones has severe resource constraints, low CPU resources and restricted memory. This makes the design and efficient implementation of text-to-speech synthesizers for these devices a challenging task. It is also emphasized that this technology may help visually and vocally impaired people. A case study of a speaking aid mobile phone and the first automatic SMS-reading mobile phone application introduced in Hungary in October 2003 are presented.

In Chapter 7, Harry Blanchard and Steve Lewis update their former treatment of design for voice mail applications which appeared in the first edition. The title of this chapter is “Voice Messaging User Interface”. Since 1999, voice mail continues to be a common application deserving special attention. Automatic speech recognition and speech synthesis is involved in the creation of unified message solutions over the telephone. Voice, e-mail and fax messages are stored in the same mailbox. This unified mailbox is a multimedia one. The user can use a visual (computer) user interface or an audio one (phone).

Matt Yuschik continues the focus on voice mail in Chapter 8 in, “Silence Locations and Durations in Dialog Management”. This discussion addresses the use of silence (pause) in discourse as a mechanism to provide turn-taking cues to the user of a voice-activated service. The location and duration of silence are defined and tested for voice-controlled voice mail. A set of duration values have been identified in the frame of multilingual research which match the behavioral flow of normal human discourse in the USA, United Kingdom, Italy and Denmark. The model is applicable to other voice-activated services which accept conversational speech as input.

Nicole Yankelovich continues the natural dialog studies in Chapter 9 in her contribution “Using Natural Dialogues as the Basis for Speech Interface Design”. These studies help to uncover common patterns of behavior and can reveal unexpected insights that significantly improve application design. Four case studies are described in detail. The results show, that natural dialogs can serve as an effective starting point for a speech user interface design.

From this point, the book contains unique chapters showing different topics in the field of human factors in speech research and technology.

In Chapter 10, Dimitri Kanevsky presents “Telematics: Artificial Passenger and Beyond.” This interesting approach introduces a human-machine interface for in-vehicle technology that is based on conversational interactivity. The Artificial Passenger prevents the driver from falling asleep. The system talks to the driver and waits for the answer. Recognition of driver states like tiredness, sleepiness, anger and stress are important in predicting possible driver actions and determining how to affect driver states in a positive way. A user state model and its adaptation to the appropriate driver are described. The heart of the system is an embedded speech recognizer for in-car use.

Chapter 11 shows us another important topic of speech output technology by Michael Divay: “A Language to Write Letter-to-Sound Rules for English and French”. The pronunciation of speech sounds is highly dependent upon context. Many examples are provided about the gap between written text and the spoken form. The author presents a re-writing formalism for grapheme-to-phoneme transcription. Difficulties and solutions for both languages are presented. In the appendices, the sounds of French and English are presented together with grapheme and sound frequencies.

A psycholinguistic topic involving research of spontaneous speech is presented in Chapter 12 by Mária Gósy and Magdolna Kovács: “Virtual Sentences of Spontaneous Speech: Boundary Effects of Syntactic-Semantic-Prosodic Properties.” The authors used spontaneous speech and its filtered form in their experiment. Filtering made the speech incomprehensible, except for suprasegmental features. The aim of the experiment was to demonstrate people’s ability to identify sentence boundaries when listening to spontaneous speech. The results indicated that boundaries are perceived primarily by pause and syntactic structure and secondarily by F0, lengthening, semantic features and other parameters. The significance of these results for speech recognition is also discussed.

Chapter 13 gives us another look into a speech generation algorithm by Michael Divay and Ed Bruckert, entitled “Text-to-Speech Formant Synthesis for French.” Here, we see the analysis of the successive components of a formant-based solution. Although there are other approaches available in designing a text-to-speech synthesizer from the acoustic point of view, formant synthesizers still have the advantages of being small in computer memory size and easy to modify to get new voice characteristics. The signal generation takes place in a hybrid synthesizer, and both cascade and parallel modules are applied.

In Chapter 14, John Thomas, Sara Basson and Daryle Gardner-Bonneau update their chapter from the first edition. The title is now “Accessibility and Speech Technology: Advancing Toward Universal Access.” The concept of universal design implies designing applications that are usable by everyone, including people with disabilities. Current assistive technology applications employing speech technology, human-computer

interaction, and the role of technical standards in accessibility are discussed. The authors mention that technological progress can be a step backward for special populations when accessibility is not considered. The role of standardization should not be underestimated. If the researchers continue to collect the data they need and develop best practices and standards, we may achieve the goal of universal access for services.

The last chapter, Chapter 15, shows us an unusual application of synthetic speech, with Mária Gósy's "Synthesized Speech Used for the Evaluation of Children's Hearing and Speech Perception." There are two main points from this research. The first is that synthetic speech can be manipulated and acoustic redundancy can be eliminated from the speech wave. The second is that manipulated synthetic speech can be used for the detection of hearing loss in children. The measurement method has already been used during hearing screenings for many years in Hungary. This chapter is a reminder that there are many potential applications of speech technology that can promote the health and welfare of the public.

Hopefully, the contributions in this second edition will convince the readers that human factors in speech technology are of continued importance in building more effective and natural solutions for human-machine interaction.

MEETINGS, CONFERENCES AND WORKSHOPS

2008

11–12 January 2008

Budapest Uralic Workshop 6 (BUW 6)

Budapest (Hungary)

<http://www.nytud.hu/bum6> (from 20 Aug 2007)

sipos@nytud.hu

11–14 January, 2008

The 6th Annual Hawaii International Conference on Arts and Humanities

Honolulu, Hawaii (USA)

<http://www.hichumanities.org/>

humanities@hichumanities.org

16–17 January, 2008

1st National Seminar on Speech and Language Disorders: Assessment and Intervention in the Indian context

Berhampur, India

gouriraj@sancharnet.in

17–18 January, 2008

CUNY Phonology Forum Conference on the Syllable

New York (USA)

syllable@cunyphonologyforum.net

18 January, 2008

The 2nd Czecho-Slovak Conference of the International Society of Phonetic Sciences

Prague (Czech Republic)

dubeda@ff.cuni.cz

23 January, 2008

Corpora in Phonological Research

Toulouse (France)

anne.przewozny@univ-tlse2.fr

24–26 January, 2008

Old World Conference in Phonology 5

Toulouse (France)

anne.przewozny@univ-tlse2.fr

9–10 February, 2008

The 1st Nordic Conference of Clinical Linguistics (NorConfClinLing2008)

Joensuu (Finland)

<http://cc.joensuu.fi/linguistics/NorConfClinLing2008/>

NorClinLing2008@joensuu.fi

18–21 February 2008

Le Changement Linguistique et ses Théories (ED-M3-2008)

Fribourg (Switzerland)

http://www2.unine.ch/structuration_periodes

mathieu.avanzi@unine.ch

- 22–23 February, 2008
Current Approaches to Spanish & Portuguese Second Language Phonology
 Minneapolis (USA)
<http://spanport.cla.umn.edu/L2phonology>
face002@umn.edu
- 22–24 February, 2008
Penn Linguistics Colloquium (PLC 32)
 Philadelphia (USA)
<http://www.ling.upenn.edu/Events/PLC/plc32@ling.upenn.edu>
- 27 February, 2008
Methodological Aspects of Intonation Research
 Bamberg (Germany)
pia.bergmann@germanistik.uni-freiburg.de
- 27–29 February, 2008
A Comparison of Signed and Spoken Languages
 Bamberg (Germany)
wrobel@daf.uni-muenchen.de
- 28–29 February, 2008
LangTech 2008 Conference
 Rome (Italy)
<http://www.langtech.it/en/secretariat@langtech.it>
- 28–29 February, 2008
The Role of Phonology in Reading Acquisition
 Bamberg (Germany)
kathrin_schrader@gmx.de
- 6–8 March, 2008
English as a Lingua Franca (ELF Forum)
 Helsinki (Finland)
<http://www.eng.helsinki.fi/ELFForum/ELF-Forum@helsinki.fi>
- 7–8 March, 2008
African American Women's Language Conference '08
 San Antonio, Texas (USA)
sonja.lanehart@utsa.edu
- 25 March, 2008
Categorical Phonology and Gradient Facts (GLOW Phonology Workshop)
 Newcastle upon Tyne (UK)
glow31@ncl.ac.uk
- 27–28 March, 2008
Belgrade International Meeting of English Phoneticians (BIMEP08)
 Belgrade (Serbia)
biljana.cubrovic@gmail.com

- 27–28 March, 2008
The 6th International Conference on Informatics and Systems – Special Track On Natural Language Processing (INFOS 2008)
Cairo (Egypt)
<http://www.fci.cu.edu.eg/INFOS2008/>
- 27–28 March, 2008
3rd Northern Englishes Workshop (NEW 3)
Salford (UK)
<http://www.esri.salford.ac.uk/esri/m/?s=4>
P.Tipton@salford.ac.uk
- 28 March, 2008
Workshop on Empirical Approaches to Speech Rhythm
London (UK)
<http://www.phon.ucl.ac.uk/rhythm2008>
rhythm2008@phon.ucl.ac.uk
- 28 March, 2008
The Neurocognition of Memory and Language
Washington DC (USA)
<http://cbbc.georgetown.edu/workshops/2008RA.html>
cbbc@georgetown.edu
- 2–4 April, 2008
14. Arbeitstagung zur Gesprächsforschung
Mannheim (Germany)
<http://www.gespraechsforschung.de/tagung/programm.htm>
tagung@gespraechsforschung.de
- 2–5 April, 2008
Didactique du français par la pratique théâtrale
Tunis (Tunisia)
colloquetunis@googlemail.com
- 3–5 April, 2008
5. Interdisziplinäre Tagung über Sprachentwicklungsstörungen (ISES 5)
Mainz (Germany)
ses5@fh-fresenius.de
<http://ises5.fh-fresenius.de>
- 3–5 April, 2008
Sociolinguistic Issues in the Use of Language
Amsterdam (The Netherlands)
<http://www.taalstudio.nl>
- 4–5 April, 2008
Phonology, Syntax and the Lexicon: Interdependence (ALOES 2008)
Paris (France)
<http://www.univ-paris13.fr/CRIDAF/avril2008.htm>
nballier@free.fr

5 April, 2008

SFU Phonology Fest 2008

Burnaby (Canada)

http://www.sfu.ca/linguistics/phonfest/phonfest_index.html

kns3@sfu.ca

11–13 April, 2008

Experimental and Theoretical Advances in Prosody

Ithaca (USA)

<http://ling.cornell.edu/prosody08>

prosody08@gmail.com

18–20 April, 2008

The 8th Phonetic Conference of China (PCC2008)

Beijing (China)

http://www.pcc2008.cn/PCC2008_en/index.html

phonetics2008@gmail.com

25–26 April, 2008

The 3rd “Talking Across the World” Conference 2008 (TAW 3)

Bangalore (India)

<http://www.talkingacrosstheworld.org>

egconf@polyu.edu.hk

2–3 May, 2008

Linguistic Variation Across the Lifespan

Columbus (USA)

springsym@ling.osu.edu

<http://www.ling.ohio-state.edu/~springsym/>

5 May, 2008

First EMUS Conference – Expressivity in Music and Speech

Campinas (Brazil)

<http://recherche.ircam.fr/equipes/analyse-synthese/EMUS>

6–9 May, 2008

Speech Prosody 2008

Campinas (Brazil)

<http://www.sp2008.org/>

sp2008_info@iel.unicamp.br

5–7 May, 2008

The International Workshop on Spoken Languages Technologies for Under-Resourced Languages (SLTU)

Hanoi (Vietnam)

<http://www.mica.edu.vn/sltu>

9–11 May, 2008

Phonetics of the Slavic Languages

New Haven (USA)

odi.reich@yale.edu

09–11 May, 2008

5th North American Phonology Conference (NAPhC5)

Montreal (Canada)

<http://linguistics.concordia.ca/naphc5/>

cogsci@alcor.concordia.ca

17–18 May, 2008

The 2nd Conference on Language, Discourse & Cognition (CLDC-2)

Taipei (Taiwan)

cldc2008@ntu.edu.tw; d94142001@ntu.edu.tw

<http://homepage.ntu.edu.tw/~gilntu/>

22–24 May, 2008

16th Manchester Phonology Meeting (16MFM)

Manchester (UK)

<http://www.englang.ed.ac.uk/mfm/16mfm.html>

patrick.honeybone@ed.ac.uk

28–30 May, 2008

Language Resources and Evaluation Conference

Marrakech (Morocco)

<http://www.lrec-conf.org/lrec2008/>

31 May – 1 June, 2008

2008 Conference of Japan Second Language Acquisition (J-SLA2008)

Kyoto (Japan)

<http://www.j-sla.org/e/index.html>

shunji12@yahoo.co.jp

4–5 June, 2008

4th Intercultural Rhetoric and Discourse Conference (IR Conference)

Indianapolis (USA)

<http://www.iupui.edu/~icic/IRconference.htm>

uconnor@iupui.edu

9–13 June, 2008

XXVII^{es} Journées d'Étude sur la Parole (JEP'08)

15^e Conférence sur le Traitement Automatique des Langues Naturelles (TALN'08)

10^e Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL'08)

Avignon (France)

<http://www.lia.univ-avignon.fr/jep-taln08/>

jean-francois.bonastre@univ-avignon.fr

12–13 June, 2008

The 5th International Workshop on Natural Language Processing and Cognitive Science (NLPCS 2008)

Barcelona (Spain)

<http://www.iceis.org/workshops/nlpcs/nlpcs2008-cfp.html>

- 13–14 June, 2008
Prosody-Syntax Interface Workshop (PSI 2)
Berlin (Germany)
<http://www.zas.gwz-berlin.de/events/psi2/>
downing@zas.gwz-berlin.de
- 16–18 June, 2008
Perception and Interactive Technologies for Speech-Based Systems (PIT08)
Kloster Irsee (Germany)
<http://it.e-technik.uni-ulm.de/World/Research.DS/irsee-workshops/pit08/introduction.html>
- 17–18 June, 2008
Prosody of Expressivity in Speech and Music
Paris (France)
<http://recherche.ircam.fr/equipes/analyse-synthese/EMUS>
- 19–20 June, 2008
ACL Special Interest Group in Morphology and Phonology (SIGMORPHON 2008)
Columbus, Ohio (USA)
<http://phonology.cogsci.udel.edu/sigmorphon2008/>
sigmorphon2008@udel.edu
- 25–28 June, 2008
12th Congress of the International Clinical Phonetics and Linguistics Association
Istanbul (Turkey)
<http://www.icpla2008.org>
dilkom@anadolu.edu.tr
- 26–28 June, 2008
Structural Features of Varieties of French in Contact (PCF)
New Orleans (USA)
klingsler@tulane.edu
- 26–28 June, 2008
First SignTyp Conference
Storrs, CT (USA)
<http://linguistics.uconn.edu/sign/>
signtyp@uconn.edu
- 29 June – 4 July, 2008
Acoustics'08 Paris
Paris (France)
<http://www.acoustics08-paris.org/>
- 30 June – 4 July, 2008
Introduction à la Phonologie Déclarative
Porto (Portugal)
jveloso@letras.up.pt
- 30 June – 2 July, 2008
11th Laboratory Phonology Conference
Wellington (New Zealand)
<http://www.vuw.ac.nz/labphon11>
labphon11@vuw.ac.nz

7–8 July, 2008

The Nature and Development of L2 French
Southampton (UK)
<http://www.floc.soton.ac.uk/conferences.html>
annabelle.david@ncl.ac.uk

9 July, 2008

The Phonological Deficit Hypothesis
Asheville (USA)
<http://www.triplesr.org/conference>
braze@haskins.yale.edu

9–11 July, 2008

Rencontre « André Martinet, linguiste »
Paris (France)
f.i.dhiver@wanadoo.fr

9–12 July, 2008

Congrès mondial de linguistique française (CMLF-08)
Paris (France)
<http://www.ilf.cnrs.fr/spip.php?rubrique4>
benoit.habert@ens-lsh.fr

10–12 July, 2008.

Choice for Voice 2008
London (UK)
www.british-voice-association.com

July 12–15, 2008

International Conference on Machine Learning and Cybernetics (ICMLC 2008)
Kunming (China)
<http://www.icmlc.com/welcome.htm>

13–16 July, 2008

International Professional Communication Conference (IPCC)
Montréal (Canada)
IPCC2008@gmail.com

21–26 July, 2008

Workshop on Speech Sciences in Linguistics (CIL 18)
Seoul (South Korea)
http://www.cil18.org/workshop/workshop_04.htm
conference@speechsciences.org

28 July – 1 August, 2008

The 11th International Congress for the Study of Child Language (IASCL 2008)
Edinburgh (UK)
<http://www.in-conference.org.uk/IASCL/>
IASCL@in-conference.org.uk

31 July – 2 August, 2008

Consonant Clusters and Structural Complexity

Munich (Germany)

<http://www.phonetik.uni-muenchen.de/cluster>

cluster@phonetik.uni-muenchen.de

4–6 August, 2008

10th Nordic Prosody conference

Helsinki (Finland)

www.helsinki.fi/speechsciences/np2008/

4–7 August, 2008

Language, Communication and Cognition

Brighton (UK)

<http://www.languageandcognition.net>

LCC@Brighton.ac.uk

4–8 August, 2008

13th International Conference on Methods in Dialectology (Methods XIII)

Leeds (UK)

<http://www.leeds.ac.uk/english/methods.htm>

engmeth@leeds.ac.uk

10–16 August, 2008

VISPP Summer School 2008

Kuressaare (Estonia)

<http://vispp2008@phon.ioc.ee>

vispp2008@phon.ioc.ee

18–22 August 2008

22nd International Conference on Computational Linguistics

Manchester (UK)

<http://www.coling2008.org.uk/>

18–22 August, 2008

Summer School on Corpus Phonology

Augsburg (Germany)

<http://www.uni-augsburg.de/summerschool>

summerschool08@phil.uni-augsburg.de

25–27 August, 2008

Special Session on Greek Phonetics (ExLing2008)

Athens (Greece)

<http://www.exling.gr/2008/home2008.htm>

exling@phil.uoa.gr

1–4 September, 2008

Intensive Workshop: Intonation in English

Kowloon (Hong Kong)

<http://www.engl.polyu.edu.hk/events/intonworkshop2008/>

egclaw@inet.polyu.edu.hk

3–5 September, 2008

AFLS Conference

Oxford (UK)

<http://www.afls.net/conferences.html>

afls2008oxford@hotmail.fr

8–12 September, 2008

11th International Conference on Text, Speech and Dialogue (TSD 2008)

Brno (Czech Republic)

<http://www.tsdconference.org>

tsd2008@tsdconference.org

10–13 September, 2008

Eurosla 18

Aix-en-Provence (France)

<http://blog.univ-provence.fr/blog/eurosla18>

11–12 September, 2008

Workshop on Phonological Variation in Voicing

Amsterdam, Leiden (The Netherlands)

Marc.van.Oostendorp@Meertens.KNAW.nl

15–17 September, 2008

Third TIE Conference on Tone and Intonation (TIE3)

Lisbon (Portugal)

<http://www.fl.ul.pt/LaboratorioFonetica/TIE3/>

sonia.frota@mail.telepac.pt

18–20 September, 2008

Discourse Coherence – Text and Theory

Paris (France)

<http://www.celta.paris-sorbonne.fr/>

celta@paris-sorbonne.fr

20 September, 2008

Comparing Prosodies Grammatically

Cambridge, MA (USA)

<http://www.fas.harvard.edu/~lingdept/comparingProsodiesgrammatically.html>

22–26 September, 2008

Interspeech (ICSLP) 2008

Brisbane (Australia)

http://www.isca-speech.org/call4prop_interspicslp2008.htm

26–28 September, 2008

Laboratory Approaches to Spanish Phonology (SP4)

Austin (USA)

<http://www.utexas.edu/cola/conferences/lasp/main/>

moll@mail.utexas.edu

26–28 September, 2008

Living, Working and Studying in Vehicular Languages

Turku (Finland)

<http://www.hum.utu.fi/oppiaineet/ranskankieli/tutkimus/konferenssit/vehicular.html>
freder@utu.fi

28 September, 2008

Psycholinguistics in Teaching English as a Second Language (TESOL)

Reading (UK)

j.c.field@reading.ac.uk

2–5 October, 2008

7. Internationale Stuttgarter Stimmtage

Stuttgart (Germany)

www.gesprochenes-wort.de

6–8 October, 2008

Acoustics Week in Canada

Vancouver (Canada)

<http://www.caa-aca.ca>

mhodgson@interchange.ubc.ca

15 October, 2008

Les Universaux prosodiques

Paris (France)

<http://www.umr7023.cnrs.fr/>

mrusso@univ-paris8.fr

15–18 October, 2008

**2nd International Conference on Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions
and**

18th Czech-German Workshop on Speech Processing

Prague (Czech Republic)

<http://www.ufe.cz/events/cost2102.php>

23–27 October, 2008

Instrumental Phonology: Patterns and Variation

México City (Mexico)

<http://lef.colmex.mx>

eherrera@colmex.mx

31 October – 2 November, 2008

Boston University Conference on Language Development (BUCLD)

Boston, MA (USA)

<http://www.bu.edu/linguistics/APPLIED/BUCLD/>

langconf@bu.edu

4–6 November, 2008

Applications of Phonetics and Phonology On Arabic

Mafrq (Jordan)

<http://www.aabu.edu.jo/art/home.htm>

said@aabu.edu.jo, said19681@yahoo.com

7–9 November, 2008

Experimental Methods in Language Acquisition Research (EMLAR)

Utrecht (The Netherlands)

<http://www.let.uu.nl/emlar/>

emlar@let.uu.nl

13–16 November, 2008

50th Annual M/MLA Convention

Minneapolis (USA)

<http://www.uiowa.edu/~mmla>

smburt@ilstu.edu

24–26 November, 2008

Congress of Phonetics and Phonology

Niteroi, Rio de Janeiro (Brazil)

<http://sbfonetica.vilabol.uol.com.br>

mtmatta@terra.com.br

27–29 November, 2008

Prosodic Interface Relations (PIR 2008)

Stuttgart (Germany)

<http://www.ims.uni-stuttgart.de/veranstaltungen/pir2008>

pir2008-info@ims.uni-stuttgart.de

3–5 December, 2008

International Symposium: 30 Aniversari del Laboratori

Barcelona (Spain)

<http://ub.edu/labfon/simposiumc.htm>

dszmidt@ub.edu

4 December, 2008

Heard around the World

Brussels (Belgium)

<http://homepages.ulb.ac.be/~heard>

heard@ulb.ac.be

4–5 December, 2008

1st International Workshop on Cataloguing and Coding of Spoken Language Data

(CatCod 2008)

Orléans (France)

<http://www.catcod.org>

7–12 December 2008

E-Humanities – an Emerging Discipline

Indianapolis (USA)

<http://www.clarin.eu/>

clarin@clarin.eu

8–12 December, 2008

International Seminar on Speech Production (ISSP'2008)

Strasbourg (France)

<http://issp2008.loria.fr/>

15–16 December, 2008
2nd International Symposium on Universal Communication
Osaka (Japan)
<http://www.is-uc.org/2008/>

2009

5–7 January, 2009
Experimental Studies on Intonation: Phonetic, Phonological and Psycholinguistic Aspects of Sentence
Potsdam (Germany)
<http://www.ling.uni-potsdam.de/pip/daten/workshop.html>
gerrit@ling.uni-potsdam.de

15–17 January, 2009
Conference on the Foot in Phonology
New York (USA)
<http://www.cunyphonologyforum.net/foot.php>
foot@cunyphonologyforum.net

16–17 January, 2009
The Divison of Labour between Morphology and Phonology
Amsterdam (The Netherlands)
<http://www.uni-leipzig.de/~exponet/meet.htm>
doreengeorgi@gmx.de

21–24 January, 2009
6th Old World Conference in Phonology (OCP6)
Edinburgh (UK)
<http://www.lel.ed.ac.uk/ocp6>
patrick.honeybone@ed.ac.uk

4 March, 2009
Insertions and Deletions in Speech
Osnabrück (Germany)
insertions@zas.gwz-berlin.de

4–6 March, 2009
DGfS Workshop “Rhythm beyond the word” (DGfS-AG RBW)
Osnabrück (Germany)
ruben@ling.uni-potsdam.de

26–27 March, 2009
Workshop on Pharyngeals & Pharyngealisation
Newcastle upon Tyne (United Kingdom)
<http://www.ncl.ac.uk/linguistics/news/events/item/international-workshop-on-pharyngeals-pharyngealisation>
ghada.khattab@ncl.ac.uk

27 March, 2009
Regards croisés sur la prosodie du français
Paris (France)
<http://www2.unine.ch/conscilaprosodie/page26423.html>
mathieu.avanzi@unine.ch

9–10 April, 2009

International Conference on Prosody and Iconicity (ProsIco 2009)

Rouen (France)

<http://www.prosico2009.com>

hancilfr@yahoo.fr

23–25 April, 2009

Experimental Pragmatics Conference 2009 (XPrag 2009)

Lyon (France)

Contact: Coralie Chevallier

<http://xprag.l2c2.isc.cnrs.fr/XPrag/>

cchevallier@isc.cnrs.fr

1–3 May, 2009

3rd Brazilian Bilingual Conference

São Paulo (Brazil)

<http://www.playpen.com.br/congresso1bilingue.asp?version=english>

playpen@uol.com.br

28–30 May, 2009

21st International Conference on Foreign and Second Language Acquisition (ICFSLA 2009)

Szczyrk (Poland)

<http://uranos.cto.us.edu.pl/~icfsla/contact.htm>

szczyrkconference@op.pl

3–5 June, 2009

English Pronunciation: Issues & Practices (EPIP)

Chambéry (France)

colloque-epip@univ-savoie.fr

5 June, 2009

Nasal 2009

Montpellier (France)

<http://w3.umh.ac.be/~nasal/Workshop/Englishversion/home.html>

nasal2009@umh.ac.be

15 June 2009

Balto-Slavonic Natural Language Processing (BSNLP 2009)

Cracow (Poland)

<http://erssab.u-bordeaux3.fr/BSNLP>

17–19 June, 2009

Phonetics and Phonology in Iberia 2009 (PaPI 2009)

Las Palmas de Gran Canaria (Spain)

<http://www.congresos.ulpgc.es/papi2009/>

papi2009@ulpgc.es

18–19 June, 2009

6^{es} Journées d'Études Linguistiques (JEL'2009)

Nantes (France)

<http://www.lettres.univ-nantes.fr/liling/jel2009/>

olivier.crouzet@univ-nantes.fr

19 June, 2009

4th Workshop on Spanish within the Tones and Break Indices

Las Palmas de Gran Canaria (Spain)

<http://www.congresos.ulpgc.es/papi2009/workshop.html>

pilar.prieto@uab.cat

21–25 June 2009

13th International Conference “Speech and Computer” (SPECOM’2009)

Saint-Petersburg (Russia)

<http://www.specom.nw.ru>

24–26 June, 2009

Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL’09)

Senlis (France)

<http://www-lipn.univ-paris13.fr/taln09/index.php?conf=RECITAL>

June 24–26, 2009

16^e Conférence sur le Traitement Automatique des Langues Naturelles (TALN’09)

Senlis (France)

www-lipn.univ-paris13.fr/taln09

5–7 July 2009

16th International Conference on Digital Signal Processing (DSP 2009)

Santorini (Greece)

<http://www.dsp2009.org>

6–9 July, 2009

International Association of Forensic Linguists (IAFL)

Amsterdam (The Netherlands)

<http://iafl09.let.vu.nl/>

iafl09@let.vu.nl

7–10 July, 2009

International Workshop on Balto-Slavic Accentology 5 (IWoba 5)

Opava (Czech Republic)

roman.sukac@fpf.slu.cz

9–11 July, 2009

Multimodality of Communication in Children (MULTIMOD 2009)

Toulouse (France)

<http://w3.eccd.univ-tlse2.fr/multimod2009/>

guidetti@univ-tlse2.fr

13–16 July, 2009

2nd International Conference on Philology, Literatures and Linguistics

Athens (Greece)

www.atiner.gr/docs/Literature.htm

28–29 July, 2009

Psychocomputational Models of Human Language Acquisition (PsychoCompLA-2009)

Amsterdam (The Netherlands)

<http://www.colag.cs.hunter.cuny.edu/psychocomp/>

- 3–5 September, 2009
2nd ISCA Workshop on Speech and Language Technology in Education (SLaTE 2009)
Wroxall Abbey Estate, Warwickshire (UK)
<http://www.sigslate.org>
- 6–10 September, 2009
Interspeech 2009
Brighton (UK)
<http://www.phon.ucl.ac.uk/home/interspeech2009/>
- 9–11 September, 2009
Discourse & Prosody Interface (IDP 09)
Paris (France)
<http://idp09.linguist.univ-paris-diderot.fr>
idp09@linguist.jussieu.fr
- 10–13 September, 2009
The 8th International Conference on Auditory-Visual Speech Processing (AVSP) 2009
Norwich (UK)
<http://www.avsp2009.co.uk>
- 13–18 September, 2009
International Conference on Text, Speech and Dialogue (TSD 2009)
Plzeň (Czech Republic)
<http://www.tsdconference.org/>
- 14–16 September, 2009
Recent Advances in Natural Language Processing (RANLP-09)
Borovets (Bulgaria)
<http://www.lml.bas.bg/ranlp2009>
- 17–18 September, 2009
Workshop on Prosody and Meaning (WPM)
Barcelona (Spain)
<http://prosodia.uab.cat/prosodyandmeaning/home/index.php>
pilar.prieto@uab.cat
- 24–26 September, 2009
Gesture and Speech in Interaction (GESPIN 2009)
Poznań (Poland)
gespin2009@gmail.com
<http://www.ifa.amu.edu.pl/~gespin/>
- 29 September – 1 October, 2009
19th Czech-German Workshop on Speech Processing
Prague (Czech Republic)
<http://www.ufe.cz/events/workshop-2009.php>
- 14–16 October, 2009
Translating Beyond East and West
Prague (Czech Republic)
<http://utrl.ff.cuni.cz/Translation-Beyond-East-and-West/prague@ff.cuni.cz>

23 October, 2009

Searching Spontaneous Conversational Speech (SSCS 2009)

Beijing (China)

<http://ict.ewi.tudelft.nl/SSCS2009/>

6–8 November, 2009

Language and Technology Conference 2009 (LTC 2009)

Poznań (Poland)

vetulani@amu.edu.pl

16–18 November, 2009

8^{es} Rencontres Jeunes Chercheurs en Parole (RJCP)

Avignon (France)

<http://rjcp2009.univ-avignon.fr/>

20 November, 2009

De la perception à la compréhension d'une langue étrangère

Université de Strasbourg

moritz@umb.u-strasbg.fr

4–5 December 2009

3^{es} Journées de Phonétique Clinique

Aix-en-Provence (France)

<http://www.lpl-aix.fr/~jpc3/>

13–17 December, 2009

Automatic Speech Recognition and Understanding Workshop (ASRU2009)

Merano (Italy)

<http://www.asru2009.org/>

23–24 December, 2009

9th Conference on Language Engineering

Cairo (Egypt)

<http://www.esole.org>

2010

14–16 January, 2010

Conference on the Word in Phonology

New York, NY (USA)

<http://www.cunyphonologyforum.net/word.php>

4–6 February, 2010

Second Language Phonology (CASPSLP2010)

Gainesville, FL (USA)

<http://caspslp2010.edublogs.org/>

24–26 February, 2010

Prosodic Typology – State of the Art and Future Prospects

Berlin (Germany)

<http://www2.hu-berlin.de/dgfs/>

- 15–19 March, 2010
International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2010)
Dallas (USA)
<http://www.icassp2010.com/>
- 19–21 March, 2010
Ultrafest V
New Haven, Connecticut (USA)
<http://www.haskins.yale.edu/conferences/ultrafestV.html>
- 24–25 March, 2010
2nd Belgrade International Meeting of English Phoneticians (BIMEP 2010)
Belgrade (Serbia)
bimep.2010@gmail.com
- 22–24 April, 2010
European dyslexia conference
Bruges (Belgium)
<http://www.khbo.be/eda-khbo-dyslexiaconference>
- 1–3 May, 2010
New Sounds 2010
Poznań (Poland)
<http://ifa.amu.edu.pl/newsounds/>
- 3–5 May, 2010
The 2nd International Workshop on Spoken Languages Technologies for Under-resourced languages (SLTU'10)
Penang (Malaysia)
<http://www.mica.edu.vn/sltu-2010>
- 11–14 May, 2010
Speech Prosody
Chicago, Illinois (USA)
<http://www.isle.uiuc.edu/speechprosody2010/>
- 19–21 May, 2010
7th Conference on Language Resources and Evaluation (LREC 2010)
La Valletta (Malta)
<http://www.lrec-conf.org/lrec2010/>
- 23 May, 2010
Third International Workshop on Emotion: Corpora for Research on Emotion and Affect
La Valletta (Malta)
<http://emotion-research.net/sigs/speech-sig/emotion-workshop>
- 24–28 May, 2010
4th International Conference on Language and Automata Theory and Applications (LATA 2010)
Trier (Germany)
<http://grammars.grlmc.com/LATA2010/>

- 25–28 May, 2010
XXVIII^{es} Journées d'Étude sur la Parole (JEP 2010)
 Mons (Belgium)
<http://w3.umh.ac.be/jep2010>
- 8–9 June, 2010
The 7th International Workshop on Natural Language Processing and Cognitive Science
 Funchal, Madeira (Portugal)
<http://www.iceis.org/Workshops/nlpcs/nlpcs2010-cfp.htm>
- 1–3 July, 2010
Colloque du Réseau Français de Phonologie. In memoriam Nick Clements
 Orléans (France)
<http://forum.bdp3.com/appels-a-com-f23/appel-a-com-colloque-du-reseau-francais-de-phonologie-t813.htm>
rpf@univ-orleans.fr
- 4–8 July, 2010
International Conference on Conversation Analysis 2010 (ICCA10)
 Mannheim (Germany)
<http://www.icca10.org/start/>
- 6–9 July, 2010
17th Workshop on Logic, Language, Information and Computation (WoLLIC 2010)
 Brasília (Brazil)
<http://wollic.org/wollic2010/instructions.html>
- 8–10 July, 2010
12th Conference on Laboratory Phonology (LabPhon 12)
 Albuquerque, NM (USA)
<http://www.unm.edu/~labfon12/>
labfon12@unm.edu
- 12–15 July, 2010
2nd World Congress of French Linguistics (CMLF-2010)
 New Orleans (USA)
<http://www.ilf.cnrs.fr/>
- 12–15 July 2010
3rd Annual International Conference on Philology, Literatures and Linguistics
 Athens (Greece)
www.atiner.gr/docs/Literature.htm
- 19–22 July, 2010
17^e Conférence sur le Traitement Automatique des Langues Naturelles (TALN'10)
 Montreal (Canada)
<http://www.groupe.polymtl.ca/taln2010>
- 6–10 September 2010
11th International Conference on Text, Speech and Dialogue (TSD 2010)
 Brno (Czech Republic)
<http://www.tsdconference.org/>

16–18 September, 2010

Language Teaching in Increasingly Multilingual Environments: From Research to Practice

Warsaw (Poland)

<http://www.ils.uw.edu.pl/time.html>

20–22 September, 2010

20th Czech-German Workshop on Speech Processing

Prague (Czech Republic)

<http://www.ufe.cz/events/workshop-2009.php>

26–30 September, 2010

Interspeech 2010

Makuhari (Japan)

<http://www.interspeech2010.org/>

27 September – 1 October, 2010

Summer School “Cognitive and Physical Models of Speech Production, Speech Perception and Production-Perception Interaction”

Berlin (Germany)

<http://summerschool2010.danielpape.info/>

11–16 October, 2010

XXXIII^e Colloque international de linguistique fonctionnelle

Korfu (Greece)

www.dfti.ionio.gr/silf2010

2011

19–21 May, 2011

Quatrièmes Journées de Phonétique Clinique

Strasbourg (France)

<http://misha1.u-strasbg.fr/IPS/>

27–31 August, 2011

Interspeech 2011

Florence (Italy)

[http://www.interspeech2011.org/IS2011-\(Welcome\).html](http://www.interspeech2011.org/IS2011-(Welcome).html)

CALL FOR PAPERS

The *Phonetician* will publish peer-reviewed papers and short articles in all areas of speech science including articulatory, acoustic phonetics, speech production and perception, speech synthesis, speech technology, applied phonetics, psycholinguistics, sociophonetics, history of phonetics, etc. Contributions should primarily focus on experimental work but theoretical and methodological papers will also be considered. Papers should be original works that have not been published and are not considered for publication elsewhere.

Authors should follow the guidelines of the *Journal of Phonetics* for the preparation of their manuscripts. Manuscripts will be reviewed anonymously by two experts of the field. The title page should include the authors' names and affiliations, address, e-mail, telephone, and fax numbers. Manuscripts should include an abstract of no more than 150 words and up to four keywords. The final version of the manuscript should be sent both in .doc and in .pdf files. It is the authors' responsibility to obtain written permission to reproduce copyright material.

All kinds of manuscripts should be sent in electronic form (.doc and .pdf) to the Editor. We encourage our colleagues to send manuscripts for our newly released section entitled MA research: Introduction. MA students are invited to sum up their research in the area of phonetics answering the questions of motivation, topic, goal, and results (no more than 1,200 words).

INSTRUCTIONS FOR BOOK REVIEWS



Reviews in *The Phonetician* are dedicated to books related to phonetics and phonology. Usually the editor contacts prospective reviewers. Readers who wish to review a book mentioned in the list of "Publications Received" or any other book, should address the editor about it.

A review should begin with the author's surname and name, publication date, the book title and subtitle, publication place, publishers, ISBN numbers, price, page numbers, and other relevant information such as number of indexes, tables, or figures. The reviewer's name, surname, and address should follow "Reviewed by" in a new line.

The review should be factual and descriptive rather than interpretive, unless reviewers can relate a theory or other information to the book which could benefit our readers. Review length usually ranges between 700 and 2500 words. All reviews should be sent in electronic form to prof. Judith Rosenhouse (e-mail: swantech@013.net).

ISPhS MEMBERSHIP APPLICATION FORM

Please mail the completed form to:

Secretary General:

Prof. Dr. Mária Gósy, DSc
Secretary General's Office:
Kempelen Farkas Speech Research Laboratory
Hungarian Academy of Sciences
Benczúr u. 33
H-1068 Budapest
Hungary

I wish to become a member of the International Society of Phonetic Sciences

Title: _____ Last Name: _____ First Name: _____

Company/Institution: _____

Full mailing address: _____

Phone: _____ Fax: _____

E-mail: _____

Education degrees: _____

Area(s) of interest: _____

The Membership Fee Schedule (check one):

- | | |
|--|---------------------|
| 1. Members (Officers, Fellows, Regular) | \$ 30.00 per year |
| 2. Student Members | \$ 10.000 per year |
| 3. Emeritus Members | NO CHARGE |
| 4. Affiliate (Corporate) Members | \$ 60.000 per year |
| 5. Libraries (plus overseas airmail postage) | \$ 32.000 per year |
| 6. Sustaining Members | \$ 75.000 per year |
| 7. Sponsors | \$ 150.000 per year |
| 8. Patrons | \$ 300.000 per year |
| 9. Institutional/Instructional Members | \$ 750.000 per year |

Go online at www.isphs.org and pay your dues via PayPal using your credit card.

I have enclosed a cheque (in US \$ only), made payable to ISPhS.

Date _____ Full Signature _____

Students should provide a copy of their student card.

NEWS ON DUES

Dues: Your dues should be paid as soon as it convenient for you to do so. Please note: dues are to be paid online via PayPal or sent directly to our treasurer.

Treasurer: Unless you pay your dues online via PayPal, please send them directly to Professor Ruth Huntley Bahr, Ph.D., Dept. Comm. Sci. & Dis., 4202 E. Fowler Ave., PCD 1017, Univ. South Florida, Tampa, FL 33620-8200 USA, Tel.: +1.813.974.3182, Fax: '1.813.974.0822, e-mail: rbahr@cas.usf.edu

VISA and MASTERCARD: You now have the option to pay your ISPhS membership dues by credit card if you hold a VISA or MASTERCARD using PayPal. Please visit our website, www.isphs.org, and click on the Membership and look under Dues for the underlined phrase, "paid online via PayPal." Click on this phrase and you will be directed to a PayPal page where you can pay your dues.

The Fee Schedule:

1. Members (Officers, Fellows, Regular)	\$ 30.00 per year
2. Student Members	\$ 10.000 per year
3. Emeritus Members	NO CHARGE
4. Affiliate (Corporate) Members	\$ 60.000 per year
5. Libraries (plus overseas airmail postage)	\$ 32.000 per year
6. Sustaining Members	\$ 75.000 per year
7. Sponsors	\$ 150.000 per year
8. Patrons	\$ 300.000 per year
9. Institutional/Instructional Members	\$ 750.000 per year

Special members (categories 6–9) will receive certificates; Patrons and Institutional members will receive plaques, and Affiliate members will be permitted to appoint/elect members to the Council of Representatives (two each national groups; one each for other organizations).

Libraries: Please encourage your library to subscribe to *The Phonetician*. Library subscriptions are quite modest – and they aid us in funding our mailings to phoneticians in Third World Countries.

Life members: Based on the request of several members, the Boars of Directors has approved the following rates for **Life Membership** in ISPhS:

Age 60 or older:	\$ 150.00
Age 50–60:	\$ 250.00
Younger than 50 years:	\$ 450.00