

Identifying a language or an accent: from segments to prosody

Jacqueline Vaissière, Philippe Boula de Mareüil

▶ **To cite this version:**

Jacqueline Vaissière, Philippe Boula de Mareüil. Identifying a language or an accent: from segments to prosody. Workshop MIDL 2004, Nov 2004, Paris, Carré des Sciences, France. pp.1-4. halshs-00185539

HAL Id: halshs-00185539
https://shs.hal.science/halshs-00185539

Submitted on 6 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identifying a language or an accent: from segments to prosody

Jacqueline Vaissière[1] *& Philippe Boula de Mareüil*[2]

[1] Laboratoire de Phonétique et Phonologie, Université de la Sorbonne Nouvelle (Paris III)-CNRS
19 rue des Bernardins, 75013 Paris
Tel. : ++33 (0)1 43 26 57 17 — Fax: ++33 (0)1 43 29 70 13
Mall: Jacqueline.Vaissiere@univ-paris3.fr — http://www.cavi.univ-paris3.fr/ilpga/ED/

[2] LIMSI-CNRS
Bât. 508, BP 133, 91403 Orsay CEDEX, France
Tel : ++33 (0)1 69 85 81 19 — Fax: ++33 (0)1 69 85 80 88
Mail: Philippe.Boula.de.Mareuil@limsi.fr — http://www.limsi.fr/Individu/mareuil/

ABSTRACT

Identifying a language or an *accent* relies on both segmental and suprasegmental levels: phonotactic constraints, unit distribution and function, syllable structure, vocal gestures and settings, global rhythm, stress, intonational patterns, etc. How to disentangle the influence of segmental *vs*. suprasegmental features (i.e. the phoneme string *vs*. prosody)? The aim of this communication is to gain insight in the contribution of different linguistic levels to the perception of a language or an *accent*, from segments to prosody .

1. INTRODUCTION

Everyday life experience shows that, in some cases, it is possible to identify a language or to determine the origin of an *accent*, by recognising relevant features. To what extent are naive listeners able to identify various languages or *accents,* and how to measure it quantitatively? To answer that question, the key difficulty is that certain parameters may be distinctive to certain listeners (for example locals) but not to others.

The following questions also deserve attention. What characterises a language or an *accent* physiologically, acoustically and perceptually? Does the identification rely more on the phonemic content and on the instantiation of co-articulation phenomena, on voice quality, on filled pause and hesitation phenomena, melodic *clichés,* pitch range, pitch level at phrase boundaries, durational profiles, stress location errors or other shibboleths, which are interpreted as "foreign"? Although the question is interesting from a practical and theoretical viewpoint, scientific publications hosted only little debate over the contribution of intonation to the perception of an *accent*. Studies even often concentrate on the articulation of the sole vowels (e.g. Harrington *et al*., 1996). What is more, the combination of pitch changes and the pronunciation of certain phonemes may result in the perception of a language or an *accent*, while each cue taken separately may not. Our purpose, in the remainder of this paper, is to focus on the relative importance of prosody, to identify a dialectal variety or an *accent*. Interestingly indeed, many presentations in this workshop approach language variety identification from that angle.

2. THE ROLE OF PROSODY

The word *accent* stems from Latin: *accentus* — *ad+cantus*, which is a loan translation of Greek (*prosôidia*) Is it by chance that this term has been chosen to designate both a prosodic prominence and a way of speaking which departs from a norm? The question applies to regional, social and foreign *accents* (Léon, 1993), and Rossi (1999) seems to answer it in the negative, without proving it. In the case of a foreign *accent*, viewed as the effect of a confrontation of two phonological systems in contact — from a native language (L1) and a target one (L2) —, the mastery of an L2-like prosody may be difficult to teach, from a pedagogical viewpoint, sometimes because teachers themselves are not well aware of their prosodic systems. But also, invariants probably exist, beyond regional differences, between a Parisian and a person from Marseilles speaking English, in the rhotic pronunciation, diphthongisation, vowel reduction, and other constant articulatory habits.

Ladd (1996) for his part relates his experience of an American living in Great Britain[1]: he suggests a more important plasticity of melody, which we can observe in infants in phase of language acquisition. It has been claimed that the characteristic categories of the mother tongue prosody are extracted very early — prior to the lexicon acquisition (Konopczynski, 1991; Mehler *et al*., 1998).

As mentioned earlier, it is noteworthy that little space is usually devoted to prosody, somewhat of an

[1] Speaking about the impression of condescension or arrogance which certain polite questions may give, Ladd expresses himself in these terms (Ladd, 1996): "As an American speaker resident in Britain I have long since adopted the RP tune for use in these contexts, and now find it rather difficult to switch off. Though my segmental phonetics and phonology are still almost completely American, I have on several occasions elicited chilly reactions in North America unintentionally using the RP tune, which suggests that it is the intonation, not the overall RP package, that Americans hear as condescending". Before him, Bloomfield had been struck by the rising intonation in certain British ways of saying "thank you" (Bloomfield, 1933).

unexplored field even in dialectology. The AMPER (Atlas Multimédia Prosodique de l'Espace Roman) linguistic atlas of Romance languages is promising, though (Romano & Contini, 2000). So is the TIE network (Tone and Intonation in Europe), the objective of which is threefold: to provide a platform to coordinate European projects in the domain of prosody, to combine intonation and tones (in particular from Serbo-Croatian and languages of immigrant communities like Papiamentu) in the laboratory phonology approach, to analyse European languages and non-standard varieties (Gussenhoven in preparation). Studies especially within the framework of the IViE project (Intonational Variation in English) revealed how certain intonational patterns are specific to some English varieties (Wells, 1982; Nolan & Grabe, 1997; Grabe *et al*., 2001; Yan & Vaseghi, 2002; Fletcher *et al*.,2004). This is the case with the "high rising terminal tune" (HRT), whether it is semantic or simply realisational in Ladd's classification (Wells, 1982; Ladd, 1996). Often related to the so-called Frequency Code[2], it is common mainly in narrative texts among young females, who terminate statements like questions, as if a confirmation from the interlocutor were expected, or for checking that the latter follows. A study is also dedicated to the contribution of intonation to the impression of German *accent* in English and English *accent* in German (Jilka, 2000). Moreover, IViE has inspired a project called Dialektintonation, the first objective of which is to describe prosodic variation across regional dialects in Germany (Peters *et al*., 2002). Likewise, a series of inquiries underlined different stress patterns in *patois* from the North/North-East of France, and more generally from the whole French-speaking area (Carton, 1977; Carton *et al*., 1983; Hintze *et al*., 2000)[3]. To complete the survey, let us cite a few additional studies on:

- the intonational patterns (especially interrogative) of some Italian varieties (Endo & Bertinetto, 1996; Gili Fivela & D'Imperio, 1997; D'Imperio & House, 1997; Marotta & Sorianello, 2001);
- Peninsular/Latin American Spanish (Navarro Tomás, 1944; Garrido Almiñana 1996; Quilis & Fernández, 1992; Obediente Sosa, 1998; Mora, 1998; Beckman *et al*., 2002);
- the identification of two dialectal varieties of Scots (van Leyden & van Heuven, 2003);
- the contribution of prosody to the identification of Dutch varieties (van Bezooijen & Gooskens, 1999);
- a prosodic typology of Swedish dialects based on pitch accent (of type I or II) (Bruce *et al*., 1999);
- tonal accent system in the dialects of Japanese (Sugito, 2004);
- the discrimination between European and Brazilian Portuguese on the basis of prosody (Frota *et al*., 2002);
- the identification of Arabic dialects (Barkat *et al*., 1999; Hamdi, 2002).

The latter study (Hamdi, 2002) as well as the one on Portuguese (Frota *et al*., 2002) were carried out using procedures put forth in Ramus (1999) dissertation, investigating the proportion of vocalic intervals (%V) and the standard deviation of consonantal intervals (ΔC) to discriminate between different rhythmic classes — especially syllable-timed *vs*. stress-timed. One of the weaknesses of those parameters is their sensitivity to tempo changes — suppose an utterance begins slowly and ends quickly. To capture this problem, Grabe proposes to measure the variability in duration between two successive intervals of the same type (vocalic or consonantal), by considering a "Pairwise Variability Index" (PVI) which normalises speech rate variation (Grabe & Low, 2002). Under different speech rate conditions, the PVI approach may be more robust than the %V-ΔC approach, but it yields a continuum rather than a clear cut separation between rhythmic classes.

Other studies attempt to elucidate the role of intonation rather than rhythm. Some of these studies, in addition to those of the TIE, IViE and Dialektintonation projects, are in keeping with the guiding principles of the autosegmental-metrical model of intonation (Ladd, 1996; Jilka, 2000; D'Imperio & House, 1997; Marotta & Sorianello, 2001; Beckman *et al*., 2002). Inspired by Pierrehumbert's (1981) work used in the ToBI system (TOnes and Break Indices) (Silverman *et al*., 1992), this theoretical framework is interesting for our issue, in the sense that it assumes autonomous segmental and suprasegmental tiers.

Perceptually too, prosodic clues facilitate human language acquisition: various studies in psycho-linguistics demonstrated it, throughout a battery of tasks consisting in decoding sequences of ambiguous words or in segmenting an artificial language (Banel *et al*., 1998). On the whole, however (with maybe the exception of stress and phoneme duration), the question of the role of prosody in the perception of a (foreign) *accent* has barely been tackled, as a thorough and recently published state-of-the-art pointed out (Piske *et al*., 2001). A model such as Flege's (2003) Speech Learning Model (SLM) primarily addresses the notion of phonetic similarity, construction of new perceptual categories for segments and **phoneme**

[2] This HRT is interpreted as iconic, non-arbitrary in some paralinguistic theories based on biological metaphors, according to which small larynxes and vocal cords with fast vibration rates would be manifestations of weakness and submissiveness — e.g. (Vaissière, 1995; Gussenhoven in preparation).

[3] In dialectally marked ways of speaking, the author notices additional vowel duration in forms such as *je vais à Nâncy* [Z(↔)vE(z)anA):si] ("I'm going to Nancy"), creating a trochaic rhythm which may give the impression of a "stress shift" or "backing", even though he challenges these terms: he rather interprets it as the preparation of a rising-falling contour distributed over two syllables.

acquisition — particularly in relatively experienced L2 speakers. Kuhl's (1991) Native Language Magnet (NLM) is also based on segment-sized linguistic units. As for Best's Perceptual Assimilation Model (PAM), this model mainly describes variation of discrimination between non-native **phonemes**, depending on their phonetic goodness of fit to native categories (Best *et al*., 2001).

3. FINAL REMARKS

Prosody has long been neglected or dismissed, perhaps owing to experimental difficulties, linked to appropriate equipment problems. Speech manipulation and synthesis now allow us to sort out the role of segmental and suprasegmental features in the perception of a language or an *accent.* Speech synthesis is a good tool to make allowances, since it enables us to monitor a number of parameters: that is the reason why it has been used for delexicalisation and monotonisation purposes (Ramus & Mehler, 1999). It has been used, together with simulated or altered speech in research on foreign *accent* (Flege *et al*., 1997; Pallier *et al*., 1997; Magen, 1998; Jilka, 2000). To obliterate (most of) the segmental structure, low-pass filtered speech is often used — typically at 400 Hz, as if spoken through a thick wall. Nevertheless, this technique entails unnatural stimuli where intelligibility is suppressed, without all the segmental information contained in microprosody being removed. In addition, it is more difficult to employ with female voices, whose pitch sometimes rises higher than 400 Hz. Consequently, it is not ascertained that the manipulation eliminates any phonemic indication; and that it does not introduce distortions into the perception of prosodic events. Text-to-speech (TTS) synthesis presents fewer drawbacks and is more ecological, since it preserves comprehension.

In several studies of the present workshop, which in return may find applications in automatic language identification, TTS synthesis is regarded as a tool. But it is also of practical and theoretical interest to understand why a synthesis voice is perceived with this or that *accent*, which can be caricatured.

In this volume, several papers investigate the identification of language varieties or *accents*. To study prosody separately, independently of segmental properties, the methodology which is proposed by e.g. Gendrot, Lehka, Martin, Romano and Vieru-Dimulescu can apply to the issue of foreign, regional or social *accent* (e.g. rural *vs*. urban). But when two forms of the same language are mutually intelligible, the need to be understood is not motivated in the same way (Munro *et al*., 1999). We come up against the matter of the prestige attached to this or that variety, this or that dialect. Working on languages of comparable status enables us to factor out a whole range of stylistic, historical and social parameters — social class, levelling (Armstrong, 2003), etc.

REFERENCES

Armstrong, N. (2003) La variation en français à la lumière de la sociolinguistique, *La Tribune Internationale des Langues,* **33**, 24-30.

Banel, M.-H., Frauenfelder, U. & Perruchet, P. (1998) Contribution des indices métriques à l'apprentissage d'un langage artificiel, *Actes des 22es Journées d'Étude sur la Parole*, Martigny, 29-32.

Barkat, M., Ohala, J. & Pellegrino, F. (1999) Prosody as a Distinctive Feature for the Discrimination of Arabic Dialects, *Proceedings of the 6th European Conference on Speech Communication and Technology*, Budapest, 395-398.

Beckman, M.E., Díaz-Campos, M., Tevis McGory, J., Morgan, T.A. (2002) Intonation across Spanish, in the Tones and Break Indices framework, *Probus,* **14***,* 9-36.

Best, C.T., McRoberts, G.W. & Goodell, E. (2001) Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system, *Journal of the Acoustical Society of America,* **109**(2), 775-794.

Bloomfield, L. (1933) *Language*, Tübingen: Max Niemeyer Verlag.

Bruce, G., Elertb, C.-C., Engstrandc, O. & Wretlingb, P. (1999) Phonetics and phonology of the Swedish dialects — a project presentation and a database demonstrator, *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, 321-324.

Carton, F. (1977) Insistance dialectale, *Studia phonetica,* **12**, 59-92.

Carton, F., Rossi, M., Autesserre, D., Léon, P. (1983) *Les Accents des Français*, Paris: Hachette.

D'Imperio, M. & House, D. (1997) Perception of questions and statements in Neapolitan Italian, *Proceedings of the 5th European Conference on Speech Communication and Technology*, Rhodes, 251-254.

Endo, R. & Bertinetto, P.M. (1996) Aspetti dell'intonazione in alcune varietà dell'italiano, *Atti delle 7e Giornate di Studio del Gruppo di Fonetica Sperimentale*, Napoli, 27-49.

Flege, J.E., Bohn, O.-S. & Jang, S. (1997) Efects of experience on non-native speaker's production and perception of English vowels, *Journal of Phonetics,* **25**, 437-470.

Flege, J.E. (2003) Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and Phonology in Language Comprehension and Production*: *Differences and Similarities*, Berlin: Mouton de Gruyter.

Fletcher, J., Grabe, E. & Warren, P. (2004) Intonational variation in four dialects of English: the high rising tune. In S.-A. Jun (Ed.), *Prosodic typology and transcription — a unified approach*, Oxford: Oxford University Press.

Frota, S., Vigrio, M., Martins, F. (2002) Language Discrimination and Rhythm Classes: Evidence from Portuguese, *Proceedings of the 1st International Conference on Speech Prosody*, Aix-en-Provence, 315-318.

Garrido Almiñana, J. M. (1996) *Modelling Spanish Intonation for Text-to-Speech Applications*. Barcelona: Tesis doctoral, Universitat Autònoma de Barcelona.

Gili Fivela, B. & D'Imperio, M. (1997) Focalizzazione e 'phrasing' nel parlato letto. Analisi di due varietà di italiano, *Atti delle 8e Giornate di Studio del Gruppo di Fonetica Sperimentale*, Pisa, 55-66.

Grabe, M.E., Post, B. & Nolan, W.F. (2001) Modelling intonational variation in English: the IViE system, *Proceedings of the Prosody 2000 Workshop*, Kraków, 51-57.

Grabe, E. & Low, F.L. (2002) Durational Variability in Speech and the Rhythm Class Hypothesis. . In C. Gussenhoven & E. Warner (Eds.), *Laboratory Phonology 7* (pp. 515-546), Berlin & New York: Mouton de Gruyter.

Gussenhoven, C. (in preparation) *The Phonology of Tone and Intonation* (http://www.let.kun.nl/tie/).

Hamdi, R. (2002) Variation rythmique des dialectes arabes, *Actes des 24es Journées d'Étude sur la Parole*, Nancy, 68-71.

Harrington, J., Cox, F. & Evans, Z. (1996) An acoustic study of broad, general and cultivated Australian English vowels, *Australian Journal of Linguistics*, **17**(2), 155-184.

Hintze, M.-A., Pooley, T. & Judge, A. (Eds.) (2000) *French accents: phonological and sociolinguistic perspectives*, London: AFLS/CiLT.

Jilka, M. (2000) *The contribution of intonation to the perception of foreign accent.* Stutgart: PhD thesis, University of Stutgart.

Konopczynski, G. (1991) L'intonation de question dans le langage emergent, *Actes du 12e Congrès International des Sciences Phonétiques*, Aix-en-Provence, 174-177.

Kuhl, P.K. (1991) Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not, *Perception & Psychophysics,* **50**, 93-107.

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

Léon, P. (1993) *Précis de phonostylistique. Parole et expressivité*, Paris: Fernand Nathan.

Magen, H.S. (1998) The perception of foreign-accented speech, *Journal of Phonetics,* **26**, 381-400.

Marotta, G. & Sorianello, P. (2001) La teoria autosegmentale nell'analisi dell'intonazione interrogativa di due varietà di italiano toscano (Lucca e Siena). In F. Albano Leoni & R. Sornicola (Eds.), *Dati empirici e teorie linguistiche* (pp. 177-204), Roma: Bulzoni.

Mehler, J., Pallier, C. & Christophe, A. (1998) Language and Cognition. In M. Sabourin, F. Craik & M. Robert (Eds.), *Advances in Psychological Science. Vol. 1: Biological and Cognitive Aspects* (pp. 381-398), Hove: Psychology Press Ltd.

Mora, E. (1998) Entonación, *Español actual,* **69**, 43-50.

Munro, M., Derwing, T. & Flege, J.E. (1999) Canadians in Allabama: A perceptual study of dialect acquisition in adults, *Journal of Phonetics,* **27**, 385-403.

Navarro Tomás, T. (1944) *Manual de entonación española*, New York: Hispanic Institute.

Nolan, F. & Grabe, E. (1997) Can 'ToBI' transcribe intonational variation in British English?, *Proceedings of the ESCA Tutorial and Research Workshop on Intonation: Theory, Models and Applications*, Athens, 259-262.

Pallier, C., Bosch, L. & Sebastian Gallés, N. (1997) A limit on behavioral plasticity in speech perception, *Cognition,* **64**(3), 9-17.

Peters, J., Gilles, P., Auer, P. & Selting, M. (2002) Identification of regional varieties by intonational cues. An experimental study on Hamburg and Berlin Germans, *Language and Speech,* **45**(2), 115-139.

Pierrehumbert, J. (1981) Synthesizing intonation, *Journal of the Acoustical Society of America,* **70**(4), 985-995.

Piske, T., MacKay, I.R.A. & Flege, J.E. (2001) Factors affecting degree of foreign accent ina an L2: a review, *Journal of Phonetics,* **29**, 191-215.

Obediente Sosa, E. (1998) *Fonética y Fonología*, Mérida: Universidad de los Andes.

Quilis, A. & Fernández, J.A. (1992) *Curso de fonética y fonología españolas*, Madrid: Consejo Superior de Investigaciones Científicas.

Ramus, F. (1999) *Rythme des langues et acquisition du langage*. Paris: Thèse de doctorat, EHESS.

Ramus, F. & Mehler, J. (1999) Language identification with suprasegmental cues: A study based on speech resynthesis, *Journal of the Acoustical Society of America,* **105**(1), 512-521.

Romano, A. & Contini, M. (2000) Un progetto di Atlante geoprosodico multimediale delle varietà linguistiche romanze, *Atti delle 11e Giornate di*

Studio del Gruppo di Fonetica Sperimentale, Padova, 121-126.

Rossi, M. (1999) *L'intonation, le système du français : description et modélisation*, Paris: Ophrys.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hurschberg, J. (1992) ToBI: a standard for labeling English prosody, *Proceedings of the 1st International Conference on Spoken Language Processing*, Banff, 867-870.

Sugito, M. (2004) 50 Years of Study on Japanese Prosody, *Proceedings of the 2nd International Conference on Speech Prosody*, Nara, 15-18.

Vaissière, J. (1995) Phonetic Explanations for Cross-Linguistic Prosodic Similarities, *Phonetica,* **52**, 123-130.

van Leyden, K. & van Heuven, V.J. (2003) Prosody versus segments in the identification of Orkney and Shetland dialects, *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 1197-1200.

van Bezooijen, R. & Gooskens, C. (1999) Identification of Language Varieties. Contribution of Different Linguistic Levels, *Journal of Language and Social Psychology,* **18**(1), 31-48.

Wells, C. (1982) *Accents of English*, Cambridge: Cambridge University Press.

Yan, Q. & Vaseghi, S. (2002) A Comparative Analysis of UK and US English Accents in Recognition and Synthesis, *Proceedings of the IEEE Internatonl Conference on Acoustics, Speech and Signal Processing*, Orlando, 413-416.