19ᵉᴹᴱ COLLOQUE ANGLAIS ORAL

La qualité de voix en anglais : de la production d'un type de phonation à la perception d'une caractéristique sociale.

Voice Quality in English : from the Production of a Phonation Type to the Perception of a Social Variable.

6 & 7 avril 2018 Villetaneuse

BOOK OF ABSTRACTS

Vendredi 6 avril

09h30 : Stephan WILHELM

10h15 : Jody KREIMAN, Anita AUSZMANN & Bruce R. GERRATT

11h30 : Katharina KLUG, Christin KIRCHHUBEL,
Peter FRENCH & Paul FOULKES

12h15 : Conférencier invité : Richard WRIGHT

15h00 : Paul SARTRE

15h45 : Bruce WILEMAN

16h30 : Patricia KEATING, Jody KREIMAN & Abeer ALWAN

Voice quality in British English Its nature, functions and applications
Stephan Wilhelm

TIL Laboratory University of Burgundy CPGE Lycée Berthollet, Annecy

This presentation deals with the nature and functions of voice quality in British English and suggests that a number of practical applications can be derived from the study of the roles it plays in language.

Like e.g. Laver (1968; 1980; 1994), Sharpe (1970) or Mackenzie-Beck (2005), we take voice quality to refer to the quality of the human voice in the production of language that results from both phonatory (laryngeal) and articulatory (supralaryngeal) adjustments. This paper argues that voice quality should be included in its own right among suprasegmental elements insofar as phonatory and articulatory settings, when conceived of as components of voice quality, fulfil a function that is overlaid to that of the phonatory and articulatory features of the English phonemes.

After having listed the main types of voice quality, we argue that, in British English at least, some of them fulfil not only a linguistic or paralinguistic, but also – and most importantly – an extralinguistic function. We suggest that this could be profitably exploited for teaching English as a foreign language.

In a number of cases, there is evidence that voice quality constitutes a reliable marker of geographical origins or social status. This paper intimates that, like segmental elements or intonation innovations, specific voice quality settings may spread through processes of sociolinguistic diffusion.

Among other research fields, we suggest that all such information could prove useful not only in sociophonetics, but also in forensic linguistics, voice recognition or speech synthesis.

Bibliography :

ABERCROMBIE, D. (1967) Elements of General Phonetics. Edinburgh: Edinburgh University Press.

BUTLER, C. (1634) English Grammar. Première édition en 1633, à Oxford, par William Turner *"for the Author"*. Edition moderne par A. EICHLER (éd.) Charles Butler's English Grammar 1634, Neudrucke Frühneuenglischer Grammatiken 4, 1, Halle: Niemeyer.

CATFORD, J.C. (1964) *"Phonation Types: the Classification of some Laryngeal Components of Speech Production"*, in D. ABERCROMBIE, D. B. FRY, P. A. D. Maccarthy, N. C. SCOTT & J. L. M. TRIM (éds.), In Honour of Daniel Jones: Papers Contributed on the Occasion of His Eightieth Birthday 12 September 1961. London: Longman, 26-37.

COADOU, M. (2007) Qualité de voix et accents régionaux en anglais britannique. Thèse de doctorat non publiée. Université Aix-Marseille 1. Laboratoire Parole et Langage (LPL), CNRS : UMR 6057, Université de Provence, Aix-en-Provence.

COADOU, M. & AUDIBERT, N. (2009) *"Voice Quality and English as a Foreign Language: A Pilot Study"*, Proceedings of the 3rd International Workshop on Advanced Voice Functions Assessment (AVFA 2009). Proceedings of the 3rd International Workshop on Advanced Voice Functions Assessment (AVFA09) Madrid, Spain.

CRUTTENDEN, A. (2001a) Gimson's Pronunciation of English. London: Routledge.

DANIELS, H. (1995) *"Psycholinguistic, psycho-affective and procedural factors in the acquisition of authentic L2 phonology"*, Speak Out! Newsletter of the IATEFL Pronunciation Special Interest Group, 15, 3-10.

DI CRISTO, A. (2003) *"De la métrique et du rythme de la parole ordinaire : l'exemple du français"*, Semen 16 : <http://semen.revues.org/document2944.html> (03/10/2017).

EPSTEIN, M. A. (2002) Voice Quality and Prosody in English. Unpublished PhD thesis. University of California.

ESLING, J. H. & Wong, R. F. (1983) *"Voice Quality Settings and the Teaching of Pronunciation"*, TESOL Quarterly, 17(1), 89–95.

FOULKES, P. (2002) *"Current Trends in British Sociophonetics"*, Working Papers in Linguistics 8(3) – A Selection of Papers from NWAV 30, 75-86.

GOBL, C. & NÍ CHASAIDE, A. (2003) *"The Role of Voice Quality in Communicating Emotions, Mood and Attitude"*, Speech Communication 40, 189-212.

KOLACZYK & P. PRZEDLACKA (éds.), English Pronunciation Models: a Changing Scene. Linguistic Insights Series. Bern: Peter Lang, 311-337.

HANDLEY, Z. (2008) *"Is Text-to-Speech Synthesis Ready for Use in Computer-Assisted Language Learning?"*, Speech Communication : <http://www.sciencedirect.com/science/ar- ticle/pii/S0167639308001842> (03/10/2017)

HENTON, C. & BLAD septembre ON, A. (1988) *"Creak as a Sociophonetic Marker"*, in L. M. HYMAN and C. N. LI (éds.), Language, Speech and Mind: Studies in honour of Victoria A. New York and London: Routledge.

HONIKMAN, B. (1964) *"Articulatory Settings"*, in D. ABERCROMBIE, D. B. FRY, P. A. D. MACCARTHY, N. C. SCOTT & J. L. M. TRIM (éds.), In Honour of Daniel Jones: Papers Contributed on the Occasion of His Eightieth Birthday 12 September 1961, 73-84.

KNOWLES, G.O. (1978) *"The Nature of Phonological Variables in Scouse"*, in P TRUDGILL (éd.), Sociolinguistic Patterns in British English. London: Arnold, 80-90.

LADD, R. (2008) Intonational phonology (Cambridge Studies in Linguistics 119) (2nd ed.). Cambridge: Cambridge University Press.

LAVER, J. (1968) *"Voice Quality and Indexical Information"*, British Journal of Disorders of Communication 3, 43-54.

LAVER, J. (1980) The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press.

LAVER, J. (1994) Principles of Phonetics. Cambridge: Cambridge University Press. LEHISTE, I. (1970) Suprasegmentals. Cambridge, Massachussets: MIT Press.

LEHISTE, I. (1996) *"Suprasegmental Features of Speech"*, in N. LASS (éd.), Principles of Experimental Phonetics. St. Louis: Mosby, 226-244.

MACKENZIE-BECK, J. M. (2005) *"Perceptual Analysis of Voice Quality: the Place of Vocal Profile Analysis"*, in W. J. HARDCASTLE & J. M. MACKENZIE-BECK (eds.), A Figure of Speech: a Festschrift for John Laver. London: Laurence Erlbaum, 285-322.

MARASEK, K. (1997) EGG and Voice Quality. Unpublished PhD thesis. University of Stuttgart.

MARTINET, A. (1982) *"Linguistique, paralinguistique, extralinguistique"*, International Review of Applied Linguistics in Language Teaching, vol. 20; 1-4, 83-84.

MESSUM, P. (2007) The Role of Imitation in Learning to Pronounce. Unpublished PhD thesis. University College, London.

MESSUM, P. (2008) *"What if Children don't Learn to Pronounce by Imitation? How Should we Teach Older Learners?"*, IATEFL Pronunciation Special Interest Group Newsletter 39, 16-21.

MESSUM, P. (2010) *"Understanding and Teaching the English Articulatory Setting"*,

Speak Out! (IATEFL Pronunciation Special Interest Group Newsletter) 43, 20-29.

PELLOWE, J. & JONES, V. (1978) *"On Intonational Variability in Tyneside Speech"*, in P. TRUDGILL (éd.), Sociolinguistic Patterns in British English. London: Arnold, 101-121.

PETERS, A. (1977) *"Language Learning Strategies: does the Whole Equal the Sum of Its Parts?"* Language 53, 560-573.

ROACH, P. (1982) *"On the Distinction Between 'Stress-timed' and 'Syllable-timed' Languages"*, in D. CRYSTAL (éd.), Linguistic Controversies. London: Arnold, 73-79.

ROACH, P. (1983) English Phonetics and Phonology (2nd éd). Cambridge: Cambridge University Press.

SCHERER, K., LADD, D. & SILVERMAN, K. (1984) *"Vocal Cues to Speaker Affect: Testing Two Models"*, Journal of the Acoustical Society of America 76, 1346-1356.

SHARPE, M. C. (1970) *"Voice Quality: A suggested Framework for Description and some observations"*, in S. A. WURM & D. C. LAYCOCK (eds.), Pacific Linguistic Studies in Honour of Arthur Capell. Pacific Linguistics Series C, 13, 115-132.

WELLS, J.C. (1982) Accents of English. Vol 1: An Introduction. Cambridge: Cambridge University Press.

WILHELM, S. (2011) Innovations segmentales et suprasegmentales dans le NW Yorkshire : Implications pour l'étude du changement accentuel dans l'anglais des îles Britanniques. Thèse de doctorat. University of Burgundy: <https://tel.archives-ouvertes. fr/tel- 00665606/document> (03/10/2017).

WILHELM, S. (2012) *"Prosodie et correction phonétique"*. La Clé des Langues (Lyon: ENS LYON/DGESCO): <http://cle.ens-lyon.fr/phonetique-phonologie/prosodie-et-correction- phonetique-partie-1--171236.kjsp> (03/10/2017).

WILLEMS, N. (1982) English Intonation from a Dutch Point of View. Dordrecht : Foris Publications.

WILSON, I. L. (2006) *"Articulatory Settings of French and English Monolingual and Bilingual Speakers"*. Unpublished PhD thesis. University of British Columbia: <http:// citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.130.2207&rep=rep1&type=pdf> (03/10/2017).

YATES, K. (2003) *"Teaching Linguistic Mimicry To Improve Second Language Pronunciation"*. Unpublished MA dissertation. University of North Texas. Denton, Texas. UNT Digital Library: <https://digital.library.unt.edu/ark:/67531/meta- dc4164/ m2/1/high_res_d/thesis.pdf> (03/10/2017).

What does it mean for a voice to sound "normal"?

Jody Kreiman,[1,2] Anita Auszmann,[2] and Bruce R. Gerratt[1]

[1] Department of Head and Neck Surgery, UCLA School of Medicine, 31-19 Rehab Center, 1000 Veteran Avenue, Los Angeles, CA 90095-1794 USA

[2] Department of Linguistics, UCLA, 3125 Campbell Hall, Box 951543, Los Angeles, CA 90095-1543 Email: jkreiman@ucla.edu Telephone: +1 310-980-4695

It is unclear what is meant by "normal" voice quality, just as it is often unclear what is meant by "voice quality" in general. The notion of "normal voice" has been used refer to the manner of sound production, regardless of whether or not the feature in question has perceptible consequences; it can refer to how the voice is perceived by a listener who has particular tastes or cultural standards, with or without the presence of a physical disorder; or it can refer to the manner in which the voice is used (i.e., shouting instead of whispering during a concert or whispering instead of shouting in a crowded room could be considered abnormal). In clinical practice, the perception of abnormality is often what draws patients into voice treatment, and restoration of normal quality is often a major goal of treatment. A clearer understanding of what listeners perceive as normal and what strikes them as disordered would thus benefit clinical practice and would inform the study of voice quality in general.

To shed light on this matter, listeners (UCLA students) heard 1-sec sustained vowels produced by 200 speakers (100 male and 100 female), half of whom were recorded in the clinic (ranging from mild to fairly severe pathology) and half of whom were UCLA students with no known vocal disorder. Male and female voices were presented in separate experiments. Use of sustained vowels and controlled listening conditions were intended to help minimize variability in judgments due to the manner in which the voice was used; and by drawing speakers and listeners from the same general population, we hoped to minimize variability due to diverse cultural expectations. Listeners compared 20 voices at a time in a series of trials during which they ordered the voices along a line according to the severity of perceived vocal pathology. Any voices perceived as normal were placed in a box at one end of the line.

Analyses are underway. Preliminary results for female speakers indicate that abnormal voices were more often heard as abnormal than were normal voices, but listener agreement overall was very poor. Discussion will focus on the factors underlying the likelihood that a given voice will be heard as normal or not normal, on how normal each voice sounds, and on differences between male and female speakers. [Supported by NIH and NSF.]

Analysing voice quality as a component of forensic speaker comparison

Katharina Klug[1], Christin Kirchhübel[1,2], Peter French[1,2] and Paul Foulkes[1]

[1] Department of Language and Linguistic Science, University of York, UK

[2] J P French Associates, York, UK

{kk667 | paul.foulkes}@york.ac.uk

{christin.kirchhubel | peter.french}@jpfrench.com

Forensic phoneticians in carrying out speaker comparison tests look for features with discriminatory power in order to assist the court in determining identity/non-identity of the known suspect and the questioned speaker (French and Stevens 2013). A survey conducted by Gold and French (2011) demonstrated that most forensic phoneticians judge a speaker's voice quality (VQ) to be the most important such parameter.

Owing to its complexity, VQ is generally analysed perceptually. Acoustic information is used whenever possible to substantiate the percept (pers. comm. with forensic phoneticians at J P French Associates, Nov. 2017). Perceptual VQ analysis is impressionistic and potentially differs in the way it is interpreted between and within analysts. In order to enhance VQ assessments, this study investigates the relationship between auditory judgements and acoustic measurements.

The feature 'breathiness' has been chosen to explore the interplay between perception and acoustics. In particular, the study is interested in investigating which acoustic correlates best reflect the percept of breathiness.

Pilot data consists of three male phoneticians producing the cardinal vowels using modal voice as well as breathy voice. Commonly used measurements of breathiness are analysed: relative intensity of H1-H2 (e.g. Klatt & Klatt 1990), H1-A1 (e.g. Wayland & Jongman 2003), H1-A2 (e.g. Garellek & Keating 2011), H1-A3 (e.g. DiCanio 2009), and H2-H4 (e.g. Henton & Bladon 1985). Based on the findings of the pilot study spontaneous speech will be investigated to detect meaningful acoustic thresholds between modal voice and breathiness.

It is predicted that the results of this investigation will confirm that it is beneficial to analyse some aspects of VQ using acoustic measurements as well as auditory judgements when carrying out speaker comparison analyses. However, the challenge will be to find acoustic features that correlate reliably with the perception of 'breathiness' in poor quality forensic recordings and which can be used on spontaneous connected speech rather than 'specimen'productions of cardinal vowels.

References

DiCanio, C. T. (2009). *The phonetics of register in Takhian Thong Chong.* Journal of the International Phonetic Association, 39(2), 162-188.

French, P. and Stevens, L. (2013). *Forensic speech science.* Mark. J. Jones & Rachel-Anne Knight (eds.), The Bloomsbury companion to phonetics, 183-197.

Garellek, M. and Keating, P. (2011). *The acoustic consequences of phonation and tone interactions in Jalapa Mazatec.* Journal of the International Phonetic Association, 41(2), 185- 205.

Gold, E. and P. French. (2011). *International practices in forensic speaker comparison.* International Journal of Speech Language and the Law, 18 (2), 293-307.

Henton, C. G. and Bladon, R. A. W. (1985). *Breathiness in normal female speech: Inefficiency versus desirability.* Language & Communication, 5(3), 221-227.

Klatt, D. H. and Klatt, L. C. (1990). *Analysis, synthesis, and perception of voice quality variations among female and male talkers.* The Journal of the Acoustical Society of America, 87(2), 820-857.

Wayland, R. and Jongman, A. (2003). *Acoustic correlates of breathy and clear vowels: the case of Khmer.* Journal of Phonetics, 31(2), 181-201.

Voice quality types and uses in North American English
Richard Wright, University of Washington

Voice quality (sometimes known as phonation type, or state of the glottis), refers to the quality of vibration of the vocal folds during voiced phonation. A wide variety of voice qualities appear in the scholarly literature spanning the functional range from emotion to voice disorders to singing [1] [2] [3] [4], but linguists have traditionally broken the continuum into five basic categories based on roles they play in a language's phonology: (spread) voiceless, breathy, modal, creaky, (constricted) voiceless [5] [6]. At one end of the continuum lie the voiceless-aspirated speech sounds where the vocal folds are spread far enough apart to prevent vibration, and at the other end lies the glottal stop where there is sufficient vocal fold constriction to prevent vibration. The intermediate three categories involve quasi-periodic vibration of the vocal folds (voicing) with varying vibratory qualities. Articulation of voice quality can be modelled as an interaction between subglottal pressure, degree of vocal fold approximation (aperture), longitudinal tension of the vocal folds (stiffness), and medial thickness of the vocal folds (thickness) [7]. Relative to modal voicing, breathy voicing (sometimes referred to as murmur) is achieved with higher glottal aperture, resulting in concomitant turbulent flow, with lower vocal fold stiffness, and decreased vocal fold thickness, resulting in an acoustic signal characterized by noise, by a lower vibratory rate (perceived as lower pitch), and by increased spectral tilt [8] [9]. Creaky voicing has differing realizations depending on its linguistic role [10] [11]. Prototypical creaky voice [11] has decreased aperture, lower vocal fold stiffness, and increased vocal fold thickness, resulting a signal with an irregular and a lower vibratory rate (resulting in the perception of noise and lower pitch), and with decreased spectral tilt. Keating et al [11] identify several other types of voice quality (typically labelled 'creaky') which differ from prototypical creak on one or more dimensions including: 1) glottal fry, 2) multiply pulsed voice, and 3) nonconstricted creak (also referred to as breathy creak [1]).

In this presentation, I will focus on 'creaky voice' (broadly-defined and to the exclusion of pressed voice and breathy voice) and concentrate on its distribution in North American English. While not contrastive, creaky voice plays an important role in phonology through consonantal allophony and through prosodic marking [12] [13]. It has also been shown to contribute to the marking of a wide variety of other discourse, pragmatic, and social functions [14] [15] [16] [17] [18] [19] [20] [21]. Of recent interest is the association of voice quality with gender identity and other social-indexical features. For example, some studies have found a greater use of creaky voicing amongst male identifying speakers than in female identifying speakers [16], [17], [18], while others have found the opposite pattern [20, 21]. In research at the UW Linguistic Phonetic Laboratory using the ATAROS Corpus [22] of 16 gender-matched and 16 gender-crossed dyads, two

patterns have emerged that advocate for a more nuanced approach to treatments of creaky voice in linguistic and social contexts. The first is that vowel height effects the probability creaky voicing occurring. The second is that, controlling for well-established segmental and prosodic factors, male speakers are more likely to use creaky voicing overall; however, much of that difference derives from the fact that they are significantly more likely to use creak in male-female dyads than in male-male ones. Interestingly, female speakers in this corpus appear to be unaffected by the gender makeup of the dyad. This finding supports the social-indexical role for voice quality modulation [20] and argues for careful balancing of segmental and prosodic factors together careful control (statistical or experimental) of the social and pragmatic variables in experiments probing voice quality.

References

1. Laver, J. (1980). *The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press.*

2. Sapienza, C., Hicks, D. M., & Ruddy, B. H. (2011). *Voice disorders.* In N. B. Anderson, and G. H. Shames (Eds.), Human Communication Disorders: An Introduction (pp. 202–237). Boston: Pearson. (8th ed.).

3. Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies.* Oxford: Wiley-Blackwell.

4. Sundberg, J. (1987). *The Science of the Singing Voice.* DeKalb, IL: Northern Illinois University Press.

5. Ladefoged, P. (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago.

6. Gordon, M., & Ladefoged, P. (2001). *Phonation types: a cross-linguistic overview. Journal of Phonetics,* 29, 383–406.

7. Zhang, Z. (2016). *Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model.* Journal of the Acoustical Society of America, 139, 1493–1507.

8. Kreiman, J., & Gerratt, B. R. (2012). *Perceptual interaction of the harmonic source and noise in voice. Journal of the Acoustical Society of America,* 131, 492–500.

9. Kreiman, J., Gerratt, B. R., Garellek, M., Samlan, R., & Zhang, Z. (2014). *Toward a unified theory of voice production and perception.* Loquens, e009.

10. Garellek, M. (2012). *The timing and sequencing of coarticulated non-modal phonation in English and White Hmong.* Journal of Phonetics, 40, 152–161.

11. Keating, P., Garellek, M., and Kreiman, J. (2015). *Acoustic properties of different kinds of creaky voice, in Proceedings of the 18th International Congress of Phonetic Sciences,* Glasgow, 2015.

12. Garellek, M. (2013). *Production and perception of glottal stops.* Ph.D. thesis UCLA.

13. Zhuang, X., and Hasegawa-Johnson, M. (2008). *Towards interpretation of creakiness in Switchboard. In Proceedings of Speech Prosody* (pp. 37–40).

14. Esling, J. H., and Edmondson, J. A. (2010). *Acoustical analysis of voice quality for sociophonetic purposes.* In M. D. Paolo, & M. Yaeger-Dror (Eds.), Sociophonetics: A student's guide chapter 11. London: Routledge.

15. Gobl, C., Ní Chasaide, A. (2010). *Voice source variation and its communicative functions.* In: Hardcastle, W., Laver, J., Gibbon, F. (eds), The Handbook of Phonetic Sciences (Second Edition). Oxford: Blackwell, 378-423.

16. Henton, C. G., *Sociophonetic aspects of creaky voice, Journal of the Acoustical Society of America,* 86: S26, 1989.

17. Klatt, D. H. and Klatt, L. C. (1990). *Analysis, synthesis and perception of voice quality variations among female and male talkers,* Journal of the Acoustical Society of America, 87: 820-857.

18. Kreiman, J., Vanlancker-Sidtis, D., and Gerratt, B. R. (2005). *Perception of voice quality.* In D. Pisoni and R. Remez (eds.), Handbook of speech perception (pp. 338–362). Malden, MA: Blackwell.

19. Pépiot, E. (2014). *Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. Speech Prosody 7,* May 2014, Dublin, Ireland. pp.305-309, 2014.

20. Podesva, R., Callier, T. (2015). *Voice Quality and Identity. Annual Review of Applied Linguistics,* 35

21. Yuasa, I. P. (2010). Creaky Voice: *A New Feminine Voice Quality for Young Urban-Oriented Upwardly Mobile American Women?.* American Speech. 85 (3). 10

22. Freeman, V., Chan, J., Levow, G.-A., Wright, R., Ostendorf, M., and Zayats, V. (2014) *Manipulating stance and involvement using collaborative tasks: An exploratory comparison, in Proceedeings of Interspeech 2014.*

Voice quality in Estuary English
Paul Sartre
Université Aix-Marseille (Laboratoire Parole et Langage)
/ Université du Havre

This presentation is that of a work in progress dealing with the voice quality of (young) men living in the south of England. This project originated from the following problem – a French learner of English who is able to reproduce the segments of that language, whether Received Pronunciation or even Estuary English, and who knows the rules of prosody, will not be able to blend into the linguistic community of the south of England (nor will he be able in any other community where EE has spread). This speaker, though not necessarily identified as a Frenchman, will be tagged as foreign, because something in his voice gives him away. In order to overcome this, one needs to take into account the overall features of the voice of males from the south of England. South African phonetician Beatrice Honikman had a similar idea in the 1960s and decided to teach her French students 'articulatory settings', or the general position of supralaryngeal articulators relative to English. Results were ground-breaking – students could reach a high degree of naturalness in the target language. Every human being is indeed more or less the same when it comes to the anatomy of the phonatory apparatus (with obvious changes relative to age or gender), but the way we position our organs depends on the linguistic community you live in. English has evolved since then and it now seems that EE prevails in many parts of Britain. EE in males seems to be characterized by a special colour of voice, which is not found in other varieties – for example General American – and even less so in French. If a learner wants to sound like an EE native, he will need to learn the articulatory settings of EE. Based on audio samples of several varieties of English, this presentation will come back the main articulatory phenomena that characterize a voice and will explore possible solutions to understand what the voice quality of male EE speakers is made of.

A sociophonetic investigation of ethnolinguistic differences in voice quality among young, South African English speakers
Bruce Rory Wileman - University of Cape Town

Prior research has suggested that there may be differences in voice quality between black and white speakers of South African English who had attended well-resourced middle-class schools. The principal objective of the study is to address the question of whether there is any acoustic evidence of such differences. The study then proceeds to describe such acoustic evidence for differences in voice quality.

The author interviewed 36 female South African English speakers (18 white and 18 black) between the ages of 18 and 22. The research subjects had all attended well-resourced middle- class schools. In order to control for the possibility of substrate influences on voice quality, all black participants were of an isiXhosa language background. High quality sound recordings were conducted, consisting of both a set of read sentences as well as semi-structured interviews, the latter of which formed the core dataset for the subsequent acoustic analysis. The acoustic data were analyzed using VoiceSauce (Shue, Keating, Vicenik and Yu 2011), a program specifically designed for the acoustic analysis of voice quality. Measurements were based on automatically segmented speech samples using FAVE (Rosenfelder, Fruehwald, Evanini and Jiahong 2011) and PRAAT (Boersma and Weeninck 2015). The VoiceSauce measurement data were statistically analyzed by means of a linear mixed effects regression analysis and Wilcoxon rank sum tests using the statistical package R to evaluate the significance of ethnicity as a variable.

The effect of ethnicity was found to be significant for several measures of spectral tilt (including for example, 2K*-5K, H4*-2K*, H1*-H2* and H1*-A1*) and cepstral peak prominence with a nearly significant effect for the subharmonics-to-harmonics ratio. Black speakers exhibited consistently higher values for most harmonic differential measures (for example, H1*-A1*) overall, while white speakers exhibited higher values for fundamental frequency, harmonics-to-noise ratio and cepstral peak prominence. The author concludes that the acoustic evidence is most consistent with the hypothesis that the white speakers overall typically use a voice quality characterized by greater vocal fold constriction, thickness and stiffness in comparison to the black speakers, hypothesized to use a voice quality characterized by more breathiness. By providing a description of voice quality variation, the research contributes towards a more complete account of sociolinguistic variation in South African English.

References

Boersma, Paul and Weenink, David. 2015. *Praat: doing phonetics by computer [Computer program].* Version 5.4.17. Available from: http://www.praat.org/

Rosenfelder, Fruehwald, Evanini and Jiahong 2011. *FAVE (Forced Alignment and Vowel Extraction) Program Suite.* Available from: http://fave.ling.upenn.edu

Shue, Yen-Liang, Keating, Patricia, Vicenik, Chad and Yu, Kristine. 2011. *Voicesauce: a program for voice analysis.* Proceedings of the ICPhS XVII. 1946-1849.

The UCLA Speaker Variability Database
Patricia KEATING (Department of Linguistics, UCLA)
Jody KREIMAN (Departments of Head & Neck Surgery and Linguistics, UCLA) Abeer ALWAN (Department of Electrical Engineering, UCLA)

A new database has been collected to facilitate the study and comparison of within-speaker and between-speaker variability in voice quality in English. For the study of within-speaker variability, a variety of speech tasks in different speaking styles, over multiple recording sessions, is desirable. For the study of between-speaker variability, a large number of speakers, both men and women, is necessary. Existing databases with several speech tasks generally have relatively few speakers, while those with many speakers generally have few speech tasks. Larger corpora also often consist primarily of low-audio-quality telephone speech. Thus there is a need for a new corpus with many speakers performing several speech tasks, with high-quality audio suitable for voice analysis.

Our new database includes 105 male and 103 female undergraduate students performing a variety of speech tasks over three separate recording sessions. Speakers are a mix of monolingual, bilingual, and L2 speakers of English, but all of the recorded speech is in English. The speech tasks are:

- the vowel /a/ repeated 3 times (in all 3 sessions)
- 5 short sentences repeated 2 times (in all 3 sessions, = 30 tokens total)
- a short impromptu monolog giving instructions about something, or directions to somewhere (25-30 seconds of speech, unscripted)
- our speaker's side of a phonecall to a friend or relative (1- 2 minutes of speech, unscripted)
- talking while watching a cute pet video (1-2 minutes of speech, unscripted)
- reporting on recent conversations in which the speaker felt neutral, happy, or annoyed (1 report per session, 25-30 seconds of speech each, unscripted)

The entire corpus has been transcribed orthographically followed by forced-alignment of word-level phoneme sequences. A large subset of the corpus (170 of the 208 speakers) has permissions from the speakers for sharing their recordings with other researchers.

For voice analysis, we use our free public software VoiceSauce. VoiceSauce provides timecourse measures of many voice-related parameters, including fundamental frequency by four methods; harmonic amplitude difference measures, which can be corrected for formant effects; various cepstral harmonic-to-noise measures; as well as formant frequencies and bandwidths. In our presentation we will show how VoiceSauce makes possible efficient analysis of voice quality on a large scale, with examples from our recent work.

Samedi 7 avril

09h00 : Katsumasa SHIMIZU

09h45 : Claire PILLOT-LOISEAU, Céline HORGUES, Takeki KAMIYAMA & Sylwia SCHEUER

10h30 : Florent CHEVALIER

11h30 : Session posters :
Véronique ABASQ, Quentin DABOUIS,
Jean-Michel FOURNIER & Isabelle GIRARD
Evelyne CAUVIN
Adrien MELI & Nicolas BALLIER
Anastasija RASTOVIC, Maud PELISSIER & Emmanuel FERRAGNE

14h00 : Anita AUSZMANN, Patricia KEATING & Jody KREIMAN

14h45 : Paolo MAIRANO

15h30 : Geoff SCHWARTZ

A Cross-Language Study of VOT in the Acquisition of English Stops Produced by Asian Speakers

Katsumasa SHIMIZU

Nagoya Gakuin University

The present study examines phonetic characteristics of English stops produced by Asian language speakers, and it mainly examines VOT of the initial stops of English which these speakers produced as a second language. The languages under investigation are Japanese, Chinese (Mandarin), Korean, Thai and Burmese. Much has been done to examine phonetic characteristics of stops in Asian languages, and they are known to have different number of phonemic categories of stops: Japanese and Chinese are known to have a two-way contrast, while Korean, Thai and Burmese are known to be a three-way one. VOT has been extensively studied for a number of languages, including them, and is considered to be a main cue to distinguish the voicing categories of stops. The speakers in these languages are considered to adjust their phonation types to acquire the voicing contrast of English stops in the learning of second language (L2), and it will be significant to examine how they adjust laryngeal timing of their stop voicing to produce the voicing contrast of English stops.

The participants in the experiment are native speakers of these languages and are students studying at Japanese universities, and their English proficiency is in intermediate-level or B1 to B2 in CEFR. The number of the participants in the experiment is 8 Japanese, 12 Chinese, 12 Korean, 11 Thai and 7 Burmese speakers. Furthermore, 6 American English speakers took part in the experiment to have a reference value of English stops. Acoustic analysis of the recorded materials in these languages was made and the measurement of VOT was made by manually positioning two cursors in the display of the waveforms.

The results of acoustic measurements show that major categories of stops voicing in these languages can be distinguished by VOT values except the contrast of voiced and voiceless unaspirated ones in Burmese. Burmese is one of tone languages, and is known to have a three way contrast of voiced, voiceless unaspirated and aspirated ones. For the production of English stops, the participants showed a variation depending on the categories, and the VOT values of the speakers of different languages were compared with those of English ones. They tend to use their own values to produce voiced and voiceless stops in English, when their own are close to those of English stops. Stops in Japanese are known to have a short prevoicing for voiced stops and a short lag for voiceless ones. Japanese EFL learners showed almost the same value for English voiced stops, while they showed to increase their own values in the production of English voiceless stops. Next, Chinese speakers use their own voiceless unaspirated stops for English voiced stops, while they try to decrease their own voiceless aspirated stops to produce English voiceless stops. Further, speakers of Korean whose language has a three way contrast of stops use their own plain stops for English voiceless ones, while they try to adjust their voiceless aspirated stops to produce English voiceless ones. As for Thai speakers, they showed a long prevoicing for their voiced stops, did a short lag for their voiceless unaspirated stops, and a long lag for their voiceless aspirated stops. In producing English stops, they try to identify which category of their own stops is close to English ones, and they used their voiced stops to produce English voiced ones, while they did their voiceless aspirated stops to produce English voiceless stops. Finally, Burmese speakers seemed to use their own voiced ones to produce English voiced ones, while they used their voiceless aspirated stops to produce English voiceless ones.

As is generally known, VOT values vary with place of articulation: they increase as the place of articulation goes from bilabial stops to velar ones. There are several reasons for this tendency such as the smaller volume of the cavity after velar stops and the speed of articulators. The tendency was found in the production of English stops by these Asian EFL learners.

In the production of English stops acquired as a foreign language, the speakers showed a variation depending on VOT values in their languages, and they tend to use their own values to produce voiced and voiceless stops in English, when they identify them as those in English. For instance, Chinese speakers use their own voiceless unaspirated stops for English voiced ones, and Thai speakers apply their voiced stops to produce English voiced ones, In this way, the speakers try to identify which category in their languages is close to voiced and voiceless stops in English, and assimilate them to English ones. The process is based on the proximity of the values between the two languages, and assimilate them when they are close to those of English ones.

Selected References :

Derwing, T. M and Munro, M.J.(2015) *Pronunciation Fundamentals, John Benjamins Publishing Co.*

Docherty, G.J. 1992. *The Timing of Voicing in British English Obstruents, Foris, New York. Lisker, L. and Abramson, A.S. 1964. A cross language study of voicing in initial stops: Acoustic measurements.* Word ,20, 384-422.

Shimizu, K. 1996. A *Cross-Language Study of Voicing Contrasts of Stop Consonants in Six Asian Languages.* Tokyo : Seibido.

Etude pilote de la voix craquée chez des locuteurs anglophones et francophones en lecture dans une situation d'interaction

Claire Pillot-Loiseau[1], Céline Horgues[2], Takeki Kamiyama[3,1], Sylwia Scheuer[2]

[1]Laboratoire de Phonétique et Phonologie (LPP), UMR 7018, CNRS et Sorbonne Nouvelle Université

[2]Semantics and Syntax – Language In Action (SeSyliA), Prismes (EA 4398),
Sorbonne Nouvelle Université

[3]Linguistique empirique : Cognition, Société et Langage (LeCSeL), TransCrit
(EA 1569), Université Paris 8 Vincennes - Saint-Denis

« Série d'excitations laryngées discrètes ou d'impulsions de basses fréquences », entre 18Hz et 65Hz sans réelles différences entre hommes et femmes (Hollien et Michel, 1968), la voix craquée (creaky voice) est caractérisée par un cycle glottique irrégulier et une phase fermée plus longue qu'en voix modale (Hollien et al. 1977). Elle est perçue comme une soudaine décroissance de hauteur (Kreiman, 1982).

Si cette qualité vocale peut manifester une altération laryngée (Ylitalo et Hammarberg 1997), elle révèle aussi des contrastes phonologiques (Ladefoged et Maddieson 1996). Marqueur syntaxique (survenant surtout en fin de phrase), la voix craquée est produite par les femmes (Wolk et al. 2012), mais d'autres auteurs montrent sa présence chez les locuteurs masculins (Henton et Bladon 1987). Dépendante du dialecte, de ses localisations dans l'énoncé (Redi et al. 2001), ou d'autres facteurs environnementaux (Cantor-Cutiva et al. 2017), elle est aussi un indice des tours de parole (Ogden 2001), de l'expression de certaines émotions (Ishi et al. 2005) ou d'hésitations (Candea et al. 2005 ; Carlson et al. 2006). Présentant une fonction sociale (Wolk et al. 2012), sa prévalence en américain est avérée (Kreiman 1982 ; Abdelli-Beruh et al. 2014), alors qu'elle est beaucoup plus rare en français (Candea et al. 2005 ; Benoist-Lucy et Pillot-Loiseau, 2013).

Nous voudrions savoir au final si, dans des conversations entre étudiant.e.s anglophones et francophones, la voix craquée propre à chaque langue évolue quand lesdites langues entrent en contact (conversations en tandem linguistique). Cette étude pilote en lecture (The North Wind and the Sun/ La Bise et le soleil lu une fois dans chaque langue) par 20 paires de locuteurs – chaque paire étant constituée d'un.e anglophone et d'un.e francophone – en tandem anglais / français nous permet d'abord de savoir si les francophones et anglophones modifient leur voix craquée en fonction de la langue parlée (L1 / L2).

Après leur annotation sur la base perceptive (Kreiman 1982) et des caractéristiques acoustiques visibles sur le spectrogramme et le signal (Ishi et al. 2005), la détermination de la fréquence et le calcul de la durée relative (rapport à la durée du texte entier) des occurrences de voix craquée ont été effectués pour chaque locuteur et chaque langue.

Nous avons actuellement analysé les productions lues de 17 locutrices (9 femmes francophones en licence d'anglais (moyenne d'âge 18,9 ans, écart-type 1,5) dont la seconde langue apprise est l'anglais pour 6 d'entre elles; 8 anglophones dont une irlandaise, 4 américaines, une costaricaine, 2 britanniques (moyenne d'âge 20,6 ans, écart-type 1,4), dont le français est la deuxième langue apprise pour 6 d'entre elles) et un locuteur américain de 20 ans.

Nos premiers résultats montrent que les occurrences de voix craquée sont plus fréquentes chez les anglophones dans les deux langues d'une part, et en anglais par rapport au français chez les francophones d'autre part. On observe toutefois une importante variabilité inter-sujet, y compris entre les deux femmes britanniques. Le sujet masculin ne présente pas de différence dans la fréquence et la durée de ces occurrences avec les autres sujets.

De plus, tous sujets confondus, les durées relatives des occurrences de voix craquée ne sont pas différentes d'une langue à l'autre. Cependant, pour chaque anglophone, ces occurrences sont plus courtes en français qu'en anglais.

Ces occurrences de voix craquée peuvent ne concerner qu'un segment, comme par exemple la glottalisation (glottal stop) de /t/ (Ex. : « shone out warmly »), ou avant une consonne sonante (Ex. : « around » où la voyelle précédant /n/ est glottalisée sous la forme de voix craquée). Garellek (2015) a mis en évidence de telles glottalisations dans ces contextes phonétiques en anglais américain.

Ces occurrences de voix craquée concernent également, aussi bien en français qu'en anglais, la fin des mots prosodiques et des groupes intonatifs de notre texte, avec une longueur de ces occurrences alors plus importantes dans ce dernier cas.

En français se produisent également des glottalisations avec des occurrences de voix craquée en situation V# #V (hiatus entre une voyelle finale de mot et la voyelle initiale du mo suivant, normalement enchaîné en français), surtout chez les anglophones.

Cette étude est un premier pas vers la recherche de l'effet de l'interaction sur la qualité vocale, notamment les occurrences de voix craquée. Pour ce faire, nous projetons de compléter ces analyses de productions contrôlées par celles des productions spontanées de nos sujets.

References

Abdelli-Beruh N, Wolk L, Slavin D. (2014). *Prevalence of Vocal Fry in Young Adult Male American English Speakers.* J Voice. 28(2):185-90.

Benoist-Lucy, A., Pillot-Loiseau, C. (2013). *The Influence of language and speech task upon creaky voice use among six young American women learning French,* Interspeech 2013, Lyon, 25-29 août 2013, 2395-2399.

Candea, M., Vasilescu, I., Adda-Decker, M. (2005). *Inter- and intra-language acoustic analysis of autonomous fillers.* DISS 05, Disfluency in Spontaneous Speech Workshop, Sep 2005, Aix-en-Provence, France. pp.47-52.

Cantor-Cutiva, L. C., Bottalico, P., & Hunter, E. (2017). *Factors associated with vocal fry among college students.* Logopedics Phoniatrics Vocology, 1-7.

Carlson, R., Gustafson K. and Strangert E. (2006). *Cues for hesitation in speech synthesis, Proceedings of Interspeech,* Pittsburgh, USA 1300-1303.

Garellek, M. (2015). *Perception of glottalization and phrase-final creak.* The Journal of the Acoustical Society of America, 137(2), 822-831.

Henton, C. and Bladon A. (1988). *Creak as a Sociophonetic Marker, in Language, Speech and Mind: Studies in honour of Victoria A. Fromkin, [Ed] Larry Hyman and Charles N.* Li, 3-39 London Routeledge.

Hollien, H. and Michel, J.F. (1968). *Vocal fry as a phonatory register, Journal of Speech Hearing Research,* 11: 600-604.

Hollien, H., Girard, G.T., Coleman, R.F. (1977). *Vocal fold vibratory patterns of pulse register phonation,* Folia Phoniatr (Basel), 9: 200-205.

Ishi, C.T., Ishiguro, H. and Hagita, N. (2005). *Proposal of acoustic measures for automatic detection of vocal fry, Proceedings of Interspeech,* Lisbon, Portugal, 481-484.

Kreiman, J. (1982). *Perception of sentence and paragraph boundaries in natural conversation,* J Phonetics, 10: 163–17.

Ladefoged, P., Maddieson I. (1996). *The sounds of the world's languages, Cambridge, MA: Blackwells.*

Ogden, R. (2001). *Turn-holding, turn-yielding, and laryngeal activity in Finnish talk-in- interaction,* Journal of International Phonetic Association, 31: 139-152.

Redi, L., Shattuck-Hufnagel S. (2001). *Variations in the realization of glottalization in normal speakers.* J Phonetics, 29: 407-429.

Wolk, L., Abdelli-Beruh. N.B., Slavin D. (2012). *Habitual use of vocal fry in young adult female speakers,* J Voice, 26 (3): 111-116.

Ylitalo, Y., Hammarberg, B. (1997). *Voice characteristics, effects of voice therapy, and long- term follow-up of contact granuloma patients,* J Voice, 14: 557–566.

Du genre en variation : le cas des règles de quantité vocalique en anglais écossais
Florent Chevalier (Université de Poitiers)

Les facteurs sociaux jouent un rôle primordial dans le processus de changement linguistique. L'opposition femmes/hommes a été particulièrement documentée et les travaux des sociolinguistes ont mis en exergue la propension des femmes à intégrer une nouvelle norme de prononciation plus rapidement que les hommes, qu'il s'agisse d'évolutions qualitatives ou quantitatives (Labov 2001).

Nous proposons d'étudier cette différence de comportement linguistique au travers du cas des règles de quantité vocalique en anglais écossais. Les dialectes écossais possèdent en effet leur propre système de longueurs : la Scottish Vowel Length Rule (SVLR), selon laquelle les voyelles sont brèves sauf devant une fricative voisée (Aitken 1981). Cette règle diffère du Voicing Effect (VE) en action dans les autres variétés de l'anglais, où les voyelles sont longues devant toute consonne voisée (Gimson 1972). Le VE concurrence actuellement la SVLR : la neutralisation progressive de cette dernière ainsi que son remplacement occasionnel par le VE ont été attestés à Édimbourg (Hewlett et al. 1999), à Berwick (Watt et Ingram 2000), ainsi qu'aux Shetland (Scobbie 2005). À Glasgow, l'étude de Rathcke et Stuart-Smith (2015) a démontré une nette érosion de la SVLR, sans adoption du VE. Cette étude ne s'intéressait toutefois qu'aux hommes ; nous avons donc cherché à savoir si dans la plus grande ville d'Écosse les femmes avaient complètement neutralisé la SVLR et d'avantage adopté le VE que les hommes.

Nous avons adopté un corpus de travail comparable, soit 12 femmes de classe populaire appartenant à quatre groupes d'âge différents (adolescentes et adultes enregistrées dans les décennies 1970 et 2000). Nous avons extrait puis segmenté toutes les réalisations accentuées des deux voyelles les plus sujettes à la SVLR, /i/ et /u/, que nous avons annotées selon différents facteurs linguistiques (notamment nature du segment suivant, contexte prosodique), puis soumises à une analyse statistique linéaire à effets mixtes.

La confrontation de nos résultats avec ceux de Rathcke et Stuart-Smith atteste d'une tendance comparable entre les hommes et les femmes, soit l'érosion progressive de la SVLR et l'absence du VE (Chevalier et Stuart-Smith 2016). Cependant, la neutralisation de la SVLR, bien qu'incomplète, est significativement plus avancée chez les femmes, bien qu'on note chez les témoins féminins les plus jeunes un retour de la SVLR, inexistant chez les hommes, potentiellement lié à des facteurs identitaires récents. Sur un plan plus large, nos résultats confirment également le rôle crucial du contexte prosodique (particulièrement de la position du segment dans la phrase) dans le changement phonétique.

Enfin, nous exposerons les résultats du premier pilote d'une nouvelle étude (thèse de doctorat de l'auteur) visant à identifier et quantifier le potentiel lien de causalité entre la convergence phonétique inter-locuteurs à très court terme (Speech Accommodation Theory, cf. Giles and Smith 1979) et le changement d'une norme de prononciation à long terme au sein d'une communauté linguistique. En vue de tester la validité de cette relation théorisée notamment par Trudgill (1986), la variation dynamique au cours d'une interaction est contrôlée en détail afin de déterminer les différents facteurs linguistiques (nature du phonème, contexte prosodique ...), situationnels (type d'interaction, actes discursifs ...) et sociaux (genre, âge des locuteurs ...) susceptibles de provoquer ou freiner la convergence phonétique. À terme, la direction et la vitesse de cette variation dynamique sera comparée à l'évolution en diachronie de la prononciation de l'anglais de Glasgow au cours du XXe siècle.

Bibliographie

Aitken, A. J. (1981). The Scottish Vowel Length Rule. In M. Benskin & M. L. Samuels (Eds.), *So meny people, longages and tonges: Philological essays in Scots and Mediaeval English presented to Angus McIntosh* (pp. 131-157). Edinburgh : The Middle English Dialect Project.

Chevalier, F., & Stuart-Smith, J. (2016). Gender and real time change in the Scottish Vowel Length Rule in Glasgow. *British Association of Academic Phoneticians' 2016 colloquium, Lancaster University.*

Giles, H., & Smith, Ph. (1979). Accommodation Theory: Optimal Levels of Convergence. In H. Giles & R. N. St-Clair (Eds.), *Language and Social Psychology* (pp. 45-65). Baltimore: Basil Blackwell.

Gimson, A.C. (1972). *An Introduction to the Pronunciation of English.* London : Latimer Trend & Co.

Hewlett, N., Matthews, B., & Scobbie, J. M. (1999). Vowel duration in Scottish English speaking children. *Proceedings of the XVth ICPhS,* San Francisco, 2157-2160.

Labov, W. (2001). *Principles of linguistic change.* Vol. 2 : Social factors. Oxford : Blackwell.

Rathcke, T., & Stuart-Smith, J. (2015). *On the Tail of the Scottish Vowel Length Rule in Glasgow.* Language and Speech, 58, 1-27.

Scobbie, J. M. (2005). Interspeaker variation among Shetland Islanders as the long term outcome of dialectally varied input: Speech production evidence for fine-grained linguistic plasticity. *QMU Speech Science Research Centre Working Papers, WP-2.*

Trudgill, P. (1986). Dialects in contact. Blackwell, Oxford.

Watt, D., & Ingham, C. (2000). Durational evidence of the Scottish Vowel Length Rule in Berwick English. *Leeds Working Papers in Linguistics, 8, 205-228.*

The phonology and morphology of the core of the English lexicon

Véronique Abasq[1], Quentin Dabouis[1], Isabelle Girard[2], Jean-Michel Fournier[1]

[1]Laboratoire Ligérien de Linguistique (UMR 7270) – Université de Tours

[2]Université du Littoral – Côte d'Opale

This paper aims to give the first results of a study on the core of the English lexicon. It focuses on what Schmitt & Schmitt (2014) call "high-frequency vocabulary", i.e. the 3,000 most frequent word families (word-forms clearly sharing a base, i.e. inflected forms and transparent derivatives, belong to the same family; e.g. accepting, acceptance, unacceptable).

Our goal is to study the morphological, phonological and graphophonological properties of the core of the lexicon in order to determine what structures and generalisations are predominant in that part of the vocabulary. We also test the validity of the rules put forward by Fournier (2007, 2010). This can provide valuable insights for morphological and phonological research and also for second language teaching: which generalisations should be taught in priority to L2 learners of English?

The corpus which we analysed are the 5,000 most frequent word-forms in the SUBTLEX-UK corpus (Van Heuven et al. 2014)), which represents close to 3,000 word families. These items were coded for morphological structure and main pronunciations for British English were taken from Wells (2008).

Preliminary results

Syntactic categories: Close to half of the corpus (46%) are nouns. The next major categories are verbs (22%), proper names (13%) and adjectives (13%).

Length of words: 37% are monosyllabic, 40% a dissyllabic and 16% are trisyllabic and only 8% are longer. Unsurprisingly, monomorphemic words tend to be shorter than complex words: we hardly find any cases of monomorphemic words which are longer than three syllables, and most of them are proper names (e.g. America, February, Elizabeth).

Morphology: Half of the words in the corpus (49%) contain a suffix, and that suffix is attached to a free base which present within the corpus for 70% of words. In 69% of those words, the suffix is an inflectional suffix (e.g. -s, -ing, -ed). 38% of words have no identifiable morphological structure. 6% of words contain a historical prefix. All other types of structures are residual. Note that as word frequency decreases, morphological complexity increases.

Primary stress: 39% of words obey Fournier's (2007) "Neutral Derivation Law" and preserve the stress pattern of their base. The remaining words overall follow the "Normal Stress Rule": initial stress for monosyllables and dissyllables and antepenultimate stress for longer words. Overall, we only found 3% of words which do not follow Fournier's rules of stress placement.

Graphophonology: Overall, the exception rate to known rules is 8% and around 6% of words do not follow any rule. Two rules account for close to half of the corpus: words with digraphs have free vowels and vowels followed by a consonant cluster have checked vowels. The next rule concerns 12% of the corpus: words with the structure <-VCV#> have free vowels.

Références

Fournier, J.-M. (2007). From a Latin syllable-driven stress system to a Romance versus Germanic morphology-driven dynamics: in honour of Lionel Guierre. *Language Sciences,* 29, 218–236.

Fournier, J.-M. (2010). *Manuel d'anglais oral.* Paris: Ophrys.

Schmitt, N. & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching,* 47(4), 484–503.

Van Heuven, W. V. J. Mandera, P. Keuleers, E. & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology,* (67), 1176–1190.

Wells, J. C. (2008). *Longman Pronunciation Dictionary* (3rd ed). London: Longman.

Profilage synchronique et longitudinal
pour une typologie prosodique socioprofessionnelle

Cauvin Evelyne, CLILLAC-ARP (EA 3967) Université Paris 7
/ PRCE à l'Université de Picardie Jules Verne

Si une typologie peut être établie en phonétique expérimentale par repérage de styles de voix, tels que la voix craquée ou soufflée par exemple, comme signe de reconnaissance perceptive et de catégorisation sociale, une investigation prosodique est tout aussi digne d'intérêt. Ainsi, l'objectif poursuivi par cette étude est d'œuvrer pour construire une typologie prosodique socioprofessionnelle à partir de données acoustiques et perceptives.

Prenant principalement appui sur des études pionnières de Fónagy en psycho-phonétique ([1983] 1991) et de Crystal dans la détermination de profils prosodiques ([1982] 1992), Cauvin (2017) a établi une méthodologie de catégorisation dans une optique de hiérarchisation qualitative de la prosodie native et non native en lecture avec les outils de profilage de Hildebrandt & Backhouse (2005).

La présente étude acoustique et perceptive sur le tempo, le registre, la mélodie et le rythme procède en deux phases. La première est synchronique et repose sur l'application de ce profilage prosodique à des extraits d'enregistrements vidéo de 5 juristes, 5 astrologues, 5 hypnothérapeutes et 5 enseignants, chacun en situation professionnelle. Ensuite, une seconde étude repère les mêmes variables marquant l'évolution diachronique du leadership prosodique chez les deux figures de proue que sont Margaret Thatcher (domaine politique) et Steve Jobs (cercle des entrepreneurs). Les mesures de hauteur mélodique, durée et intensité sous-tendent les variables à la source d'un profilage personnalisé et de groupe. L'étude statistique procède par plans factoriels, cercles de corrélations et dendrogrammes (Husson et al. 2016).

Bien que les caractéristiques personnelles soient surreprésentées dans des groupes aux effectifs restreints, des critères normés se dégagent des catégories et semblent être des éléments incontournables de l'arsenal que doit posséder un individu pour être crédible et reconnu dans sa sphère socioprofessionnelle, au même titre que la teneur de son message ou sa tenue vestimentaire.

Références

CAUVIN, E. (2017). *Élaboration de critères prosodiques pour une évaluation semi-automatique des apprenants francophones de l'anglais.* Thèse de doctorat de l'Université Paris Diderot - Paris 7.

CRYSTAL, D. ([1982] 1992). *Profiling Linguistic Disability.* San Diego : Singular Publishing Group.

FÓNAGY, I. ([1983] 1991). *La vive voix : Essais de psycho-phonétique.* Paris : Payot.

HILDEBRANDT, M. & BACKHOUSE, J. (eds.) (2005). D7.2: *Descriptive analysis and inventory of profiling practices. Technical report, FIDIS (Future of Identity in Information Society) deliverable.* [www.fidis.net]

HUSSON, F., LÊ, S., PAGES, J. ([2009] 2016). *Analyse de données avec* R. Rennes : Presses Universitaires de Rennes.

Categorizing learner phonetic realisations with the knn algorithm
The case of French learners with KIT and FLEECE

Adrien Méli & Nicolas Ballier (Paris Diderot, CLILLAC-ARP)

The realisations of KIT and FLEECE of Catalan learners of English have been analysed on the basis of perception tests (Boersma & Escudero 2005, Morrison, 2007). In this paper, we propose to investigate the categorisation of learner production on an acoustic basis. We present the results of experiments that try to categorize learner data on the basis of F1 and F2 realisations. 15 speakers have been analysed in a longitudinal corpus recorded at the university of X. The formants were extracted automatically and values filtered by an inhouse algorithm. We have used BDM normalisation. In other words, Raw F1 and F2 formant values were converted. (F2-F1) and (F3-F2) were computed and, in order to normalise x and y scales (as their variation might be different), we have used a z-score.

Section 1 reports our first experiments which discussed the separability of the vowels, considering them in the BDM-normalized F1 / F2 space.

Section 2 discusses the confusion matrices obtained with the results aggregated over the four sessions and summed up in Figure 1.

Figure 1 Confusion matrix of the monophthong realisations (n= 40,235) for the 15 speakers

Section 3 will show how some of the results could be implemented in CAPT system to help learners for some self diagnosis of their realisations. Using transcriptions of the same speakers reading a text in French, we have used the knn algorithm to assign a label to their realisations of the KIT and FLEECE vowels, classified either as French /i/, British /iː/ or /ɪ/ or American /iː/ or /i/.
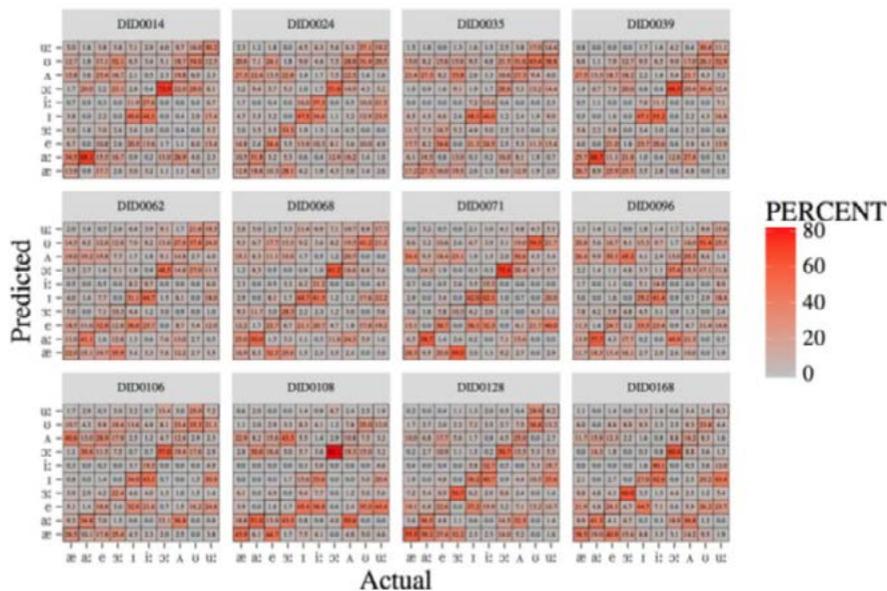
References

Boersma, P, & Escudero, P. (2005). *"Measuring Relative Cue Weighting: A reply to Morrison"*. Studies in Second Language Acquisition, 27. 607–617.

Morrison, G.S. (2007). In M. J. Solé, P. Prieto, J. Mascaró (Eds.), *Logistic regression modelling for first- and second- language perception data, Segmental and prosodic issues in Romance phonology,* 219–236, Amsterdam: John Benjamins.

The perception of swearwords by French learners of English: an experiment involving electrodermal activity

Anastasija Rastovic, Maud Pélissier, Emmanuel Ferragne CLILLAC-ARP EA 3967 – Université Paris Diderot

When people hear swearwords or taboo words, they often experience discomfort. However, it has been shown that the emotional responses caused by such words are stronger when the words are spoken in the listener's first, rather than second, language. In order to replicate these findings, we conducted a pilot perceptual experiment where French learners of English were asked to listen to English and French swearwords – emotionally neutral words were also included to serve as baseline – while their electrodermal activity (EDA) was monitored. EDA records small fluctuations in skin conductance caused by variations in the activity of sweat glands. Such variations are known to be correlated to the emotional state of the listener and, in particular, they reflect high arousal or stress.

EDA was recorded from electrodes placed on the index and middle finger of the participants' non-dominant hand while they were listening to a list of 119 words. Each stimulus was played back through headphones every 10 seconds, and the participants were asked to write down what they had heard. This task was meant to force the participants to focus on the stimuli, and it allowed the investigators to confirm that the words had been understood correctly.

For the analysis, the EDA signal was time-locked to the onset of each stimulus in an attempt to detect the expected event-related phasic skin conductance response, i.e. a sudden deflection in the EDA signal triggered by the corresponding word and occurring shortly after it. The figures below exemplify the EDA response (in normalized units) to 4 stimuli for participant VF (a native speaker of French): the first two graphs show responses induced by swearwords and the other two display responses to emotionally neutral words. The expected "hump" in the signal can be seen in Figures 1 and 2 in response to swearwords while no such pattern occurs in Figures 3 and 4. However the magnitude of the response in Figure 1 is greater than that of Figure 2 which, if this pattern were generalizable, would contradict our expectation that EDA responses to L1 swearwords should be stronger than those triggered by L2 swearwords. We are still in the process of fine-tuning the parameters of our experimental design, and therefore the final presentation will include an improved version of the pilot study we have just conducted.
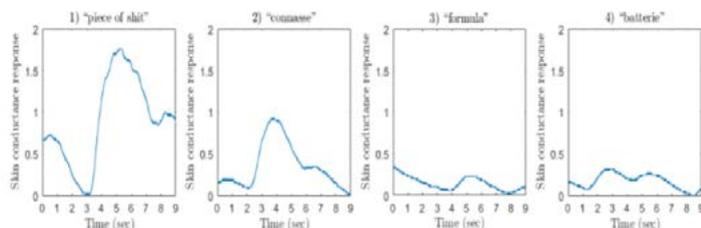
Voice quality classification in different speech styles`

Anita Auszmann[1], Patricia Keating[2], and Jody Kreiman[3]

[1.] University of California, Los Angeles; Department of Linguistics

[2.] University of California, Los Angeles; Department of Linguistics

[3.] University of California, Los Angeles; Department of Head & Neck Surgery

It is well known that different speech styles (reading, presentation, casual conversation, etc.) and conditions (talking on the phone, speaking in a noisy place, etc.) may cause changes in speech. However, it is not clear how voice quality changes under different circumstances: what kinds of modifications and acoustic changes are occurring? The present study seeks to understand how voice quality variation can be quantified on multiple parameters across different speech styles. We hypothesized that read speech is significantly different from other speech styles, given that read and spontaneous speech have been shown to differ in, e.g. intonation and rate [E. Blaauw (1991), Howell–Kadi-Hanifi (1991)]. Our database and tools made the investigation of this question possible.

In our research 70 monolingual English speakers (35 male, 35 female) from the UCLA Speaker Variability Database were investigated. We analyzed the 3 most frequent vowels across 6 different speech styles: (1) the participant had to give instructions about how to do something, (2) the participant had to speak about a recent conversation that was not important/exciting/upsetting, or that made the speaker really (3) happy, or (4) annoyed; the participant was asked to (5) read 5 randomized Harvard sentences and (6) talk to a friend or relative on their cellphone. Only condition (5) involves read speech; all others are unscripted. Extraction of acoustic features (including harmonic amplitude difference measures, harmonic- to-noise measures, fundamental frequency, formants) was carried out by VoiceSauce (Shue et al. 2011).

For statistical analysis (discriminant analysis) we used SPSS. We converted our data into a two dimensional space which is structured by functions and shows which acoustic features are doing the work in classification. The results showed that the investigated speech styles were divided into three groups which are clearly separated from each other: 1) read sentences, 2) phone call and 3) all the remaining speech styles. This classification structure is due primarily to acoustic differences in harmonic amplitudes across the frequency range, in fundamental frequency and in the strength of excitation of voicing.

Do sentiment analysis scores correlate with acoustic features of emotional speech?

Paolo Mairano, University of Turin

In this study we are concerned with the expression of emotions in speech. In particular, we investigate the possible correlation between sentiment analysis metrics and acoustic characteristics of speech as measured on English audiobooks.

[1] have highlighted that emotional and affective information in narratives can be conveyed by different components at different levels: lexical (with intensifiers, modals, hedges, etc.), syntactic (e.g. relative clauses to comment on actions and behaviours), paralinguistic (facial expressions, gestures). Additionally, abundant literature (e.g. [2], [3]) has shown that emotional speech is characterized by various acoustic cues, such as speech rate, pause duration, pitch range, pitch mean, F0 contours, voice quality parameters. However, most research on emotional speech analyses sentences produced by actors, disregarding more naturally produced speech due to the difficulty in finding suitable emotional data. We argue that audiobooks provide a convenient and slightly more ecological alternative (as already claimed by [4]), and have the advantage of being easily available online.

We therefore performed an analysis using data from 40 audiobooks, each read by a different speaker, all coming from the 'clean' (sic) section of the LibriSpeech corpus ([5]). For each syntactic period, we computed sentiment analysis scores using tools and lexicons provided by the Python-NLTK package (i.e. Vader [6] and SentiWordnet [7]) at word and sentence levels, and extracted several acoustic parameters (pitch mean, pitch range, pitch stdev, articulation rate, shimmer, jitter, HNR, etc.) via Praat scripts. In order to explore the relationship between acoustic and textual metrics, we performed a statistical analysis by building mixed-effects models with sentiment analysis scores as dependent variables and acoustic metrics as predictors. The effects of predictors tend to be small and marginally significant (similarly to what was found by [4] on data of only one audiobook), and the power of the models is usually low. We are now extending the analysis to more data, but these preliminary results seem to suggest that the different components used to express emotions (acoustic vs lexical and syntactic) tend to be complementary rather than additive, at least in read speech.

References

[1] Reilly, J., & Seibert, L. (2003). *Language and emotion.* In R.J. Davidson, K.R. Scherer and H.H. Goldsmith (ed.): Handbook of affective sciences, OUP, 535–59.

[2] Schröder, M., Cowie, R., Douglas-Cowie, E., Westerdijk, M., & Gielen, S. (2001). *Acoustic correlates of emotion dimensions in view of speech synthesis.* In 7th European Conference on Speech Communication and Technology, Aalborg (Denmark).

[3] Audibert, N., Aubergé, V., & Rilliard, A. (2005). *The prosodic dimensions of emotion in speech: the relative weights of parameters. In 9th European Conference on Speech Communication and Technology*, Lisbon (Portugal).

[4] Charfuelan, M., & Schröder, M. (2012). *Correlation analysis of sentiment analysis scores and acoustic features in audiobook narratives.* In 4th International Workshop on Corpora for Research on Emotion Sentiment & Social Signals (ES3), Istanbul (Turkey).

[5] Panayotov, V., Chen, G., Povey, D., & Khudanpur, S. (2015). *Librispeech: an ASR corpus based on public domain audio books.* In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane (Australia).

[6] Hutto, C.J. & Gilbert, E.E. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text.* 8th International Conference on Weblogs and Social Media (ICWSM-14), Oxford, MI (US).

[7] Baccianella, S., Esuli, A., & Sebastiani, F. (2010). *Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining.* In Proc. of LREC (Vol. 10, No. 2010, pp. 2200-2204).

Voice quality and L2 proficiency in the English tense-lax contrast

Geoff Schwartz (geoff@wa.amu.edu.pl) – Adam Mickiewicz University

The tense-lax contrast is one of the more notorious features of the English sound system from the point of view of second language (L2) pronunciation. Learners from countless L1 backgrounds have difficulties mastering the distinction, which is a commonly encountered and characteristic element of a foreign accent in English. Although textbooks of English pronunciation (Cruttenden 2001; Collins & Mees 2009) focus on duration, peripherality in F1-F2 space, and to some extent diphthongization, experimental studies have shown that phonation characteristics also play a role in maintaining the contrast in native varieties (Holt et al. 1994; Lotto et al. 1997). Tense vowels in English have been found to show a greater degree of breathiness, while the lax vowels exhibit a stiffer voice quality. The goal of the present paper is to investigate the degree to which highly proficient L2 speakers of English implement phonation differences in their production of the contrast.

Twenty-nine L1 Polish speakers of English, representing two levels of proficiency (B2 and C2 according to the CEFR), produced citation form productions of English vowels in a /h_d/ context. A preliminary analysis of five repetitions per speaker of the heed-hid contrast has been carried out. Voice quality was measured by comparing intensity at two different frequency bands: a low band up to 1000 Hz, and a high band between 1000 and 5000 Hz (cf. Plag et al. 2011; Sluijter & van Heuven 1996). A larger difference between the low band and high band is indicative of breathier voice quality. The intensity difference between the two bands served as the dependent variable in a linear mixed effects model with Proficiency and Tenseness as fixed factors, and Speaker as a random factor. Results showed a significant Proficiency*Tenseness interaction by which only the advanced speakers produced more breathiness in their productions of heed (Reference value: Advanced*hid = 13.34; Advanced*heed: B=1.21, Std. Error = 0.45, t=2.69, p=.008). Future work will include the who'd-hood contrast, as well as other measures of voice quality.

Implications for the phonological representation of vowel quality will also be discussed. It will be argued that phonation characteristics can contribute to the scalar realization of holistic phonological primes (something akin to elements rather than distinctive features), which allow us to better understand the evolutionary relationship between linguistic and extra-linguistic use of voice quality.

References

Collins, Beverly & Inger Mees (2009). *Practical Phonetics and Phonology, A Resource Book for Students.* London: Routledge.

Cruttenden, Alan (2001). Gimson's Pronunciation of English, 6th edition. London: Arnold. Holt, Lori, Andrew Lotto & Keith Kluender (1994). *Effect of voice quality on the tense/lax distinction for English vowels.* JASA 96. https://doi.org/10.1121/1.410935

Lotto, Andrew, Lori Holt & Keith Kluender (1997). *Effect of voice quality on perceived height of English vowels.* Phonetica 54, 76-93.

Plag, Ingo, Gero Kunter & Mareile Schramm (2011). *Acoustic correlates of primary and secondary stress in North American English.* Journal of Phonetics 39, 362-374.

Sluijter, Agaath & Vincent van Heuven (1996). *Spectral balance as an acoustic correlate of linguistic stress.* JASA 100(4), 2471-2485.