

Institution: Queen Margaret University, Edinburgh		
Unit of Assessment: UoA 26 Modern Languages and Linguistics		
Title of case study: Phonetic voice quality analysis enhances reliability of forensic voice analysis, enables commercial voice characterisation and improves vocal health		
Period when the underpinning research was undertaken: 2000-2020		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Dr Felix Schaeffler	Researcher, then Lecturer, Speech and Hearing Sciences	2006 onwards
Professor Janet Beck	Head of Division, Speech and Hearing Sciences	September 1990 onwards
Period when the claimed impact occurred: Aug 2013 - 31 Dec 2020		
Is this case study continued from a case study submitted in 2014? N		
1. Summary of the impact <p>Enhanced professional services in Forensic Phonetics:</p> <ul style="list-style-type: none"> Forensic analysis of speech data is enhanced by a specialist voice analysis approach (Vocal Profile Analysis), influencing the practice of forensic phoneticians, police, legal practitioners and their clients. <p>Commercial value in voice characterization for creative industries:</p> <ul style="list-style-type: none"> Investment created a start-up company (The Voice Distillery Ltd), exploiting our innovative approach to commercial voice-casting to streamline processes for agents, production companies and voiceover artists. <p>Support for remote voice therapy:</p> <ul style="list-style-type: none"> Our findings about which acoustic voice parameters are captured reliably with mobile phones inform international guidelines for clinical voice assessment and therapy during Covid19, with benefits for clinicians and patients. 		
2. Underpinning research <p>The multi-stranded research underpinning this impact is focussed on improving our understanding of the nature, causes and effects of voice variation (within and between speakers) and developing and identifying reliable indicators of voice quality characteristics. Our research is interdisciplinary and differs from much previous voice research by employing a theoretically informed perceptual model (Vocal Profile Analysis, developed initially by Laver and Beck and refined over several decades) as a link between the acoustic signal and qualitative interpretation in voice health and other contexts. Linguistics and phonetics provide theoretical frameworks for understanding the linguistic, paralinguistic and extralinguistic functions of voice, speech science informs acoustic analysis, and speech and language therapy provides theoretical bases for understanding the links between voice quality and vocal health.</p> <p>Research that is aimed at development, evaluation and application of specialist perceptual and acoustic techniques for voice analysis spans several decades. The work of Beck (aka Mackenzie Beck) has continued to be crucial for the development of the Vocal Profile Analysis (VPA) scheme, a phonetic tool for perceptual analysis of voice quality, from its inception. VPA involves expert auditory-perceptual judgment of the various parameters ("settings") that contribute to overall voice quality. Each setting has defined physiological and acoustic correlates, so the completed analysis provides a comprehensive indication of a speaker's habitual voice production tendencies. An</p>		

innovative aspect was the inclusion of both laryngeal factors (e.g. whisperiness, creakiness) and vocal tract adjustments (e.g. tongue or jaw settings), reflecting the complex interrelationships within the voice production system. Since 2000, research has involved refinement and evaluation of VPA (e.g. [1]) and establishment of sociophonetic voice norms needed for clinical and other applications (e.g. [2] elucidates geographical and gender-related voice variation in adolescents). Our VPA research has been a reference point for researchers in the UK, Brazil, Spain, China, Ireland and Sweden, working in sociophonetics, speech pathology, articulatory phonetics (e.g. MRI investigation of physiological correlates of settings) and forensic science.

One recurring research interest is the extraordinary range of meaningful variability in the human voice and the associated immense semantic space of voice descriptors in everyday as well as literary language (e.g. describing the sound, character attributions or interlocutor effects of voices). In [3] we argue for an intermediate phonetic level of description when mapping acoustic voice signals to voice descriptors in order to provide better descriptive structure and transparency.

VPA is the only phonetic description system that is detailed enough to serve as this intermediate level of description. Our model requires a detailed understanding of the relation between acoustic and VPA parameters. In [3] we lay the foundations of a systematic mapping of acoustics to VPA by presenting significant correlations between a subset of VPA settings and acoustic parameters. These findings and the associated ideas about levels of description have informed the foundational concept of a new commercial enterprise, as outlined below.

Acoustic analysis is an important component of clinical voice analysis, and much of our recent research focusses on the challenges associated with this, by elucidating the acoustic correlates of perceptual judgments (as in [3] above) and assessing the reliability of acoustic parameters across different environments. The potential of smartphones for health-related data collection is widely recognized, but their use in voice health depends on reliability of acoustic data. Our interdisciplinary research confirmed the reliability of smartphone recordings for certain acoustic parameters (mean F0 and cepstral peak prominence showed acceptable random error size) but not others (jitter and shimmer). It also highlighted some important methodological issues that inform clinical assessment; these have proved to be of particular relevance in the context of remote working associated with Covid-19 [4,5].

3. References to the research

Names **in bold** were QMU members of staff at the time of publication. Note: Beck and Mackenzie Beck are the same person.

- [1] **Mackenzie Beck, J.** (2005) [Perceptual analysis of voice quality: the place of Vocal Profile Analysis](#). In Hardcastle, W. & Mackenzie Beck, J. (Eds) [A Figure of Speech: a Festschrift for John Laver](#). Mahwah, New Jersey: Lawrence Erlbaum Associates, pp. 285-322. (Also publishes as Mackenzie Beck).
- [2] **Beck, J., & Schaeffler, F.** (2015) [Voice quality variation in Scottish adolescents: gender versus geography](#). *Proceedings of the 18th ICPhS, Glasgow*.
- [3] **Schaeffler, F., Eichner, M., & Beck, J.** (2019). [Towards ordinal classification of voice quality features with acoustic parameters](#). *Studenttexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2019*, 288-295.
- [4] **Jannetts, S., Schaeffler, F., Beck, J. and Cowen, S.** (2019) [Assessing voice health using smartphones: bias and random error of acoustic voice parameters captured by different smartphone types](#). *International Journal of Language and Communication Disorders* Vol 54(2), 292-305.
- [5] **Schaeffler, F., Jannetts, S. and Beck, J.** (2019). [Reliability of clinical voice parameters captured with smartphones—measurements of added noise and spectral tilt](#). In *Proceedings of the 20th Annual Conference of the International Speech Communication Association INTERSPEECH, Graz, Austria, 15-19 September 2019*. ISCA.

Evidence of 2 quality: Papers [2]-[5] are peer reviewed; [3] (0A/02/26) and [4] (0A/01/26) are submitted outputs. The quality of research in papers [1] and [2] is evidenced by widespread*

adoption by research groups and practitioners from a range of disciplines including MRI investigations of voice variation, voice therapy, expression of emotion, sociophonetics and forensic phonetics.

4. Details of the impact

1. International impact on professional services and professional education: voice analysis in Forensic Phonetics

In evaluating the probability that two or more voice samples are from the same speaker, forensic phoneticians use a multistranded suite of acoustic and perceptual analysis tools. Where acoustic quality of audio recordings used as evidence in court is inadequate for reliable acoustic analysis, perceptual tools (e.g. VPA) become especially important. VPA training based on our research [1] delivered to three key groups of forensic phoneticians as a pathway to impact (JP French Associates (JPFA) + University of York (UoY); Pontifícia Universidade Católica de São Paulo (PUC-SP) + Brazilian Federal Police; and the Netherlands Forensics Institute) generated impact in several ways.

JP French Associates, one of the world's leading forensic speech laboratories, first adopted our VPA framework in 2010 and continued to do so through the census period, creating a streamlined version specifically for forensic purposes [s1]; this version is now (2020) used by at least two other UK forensic experts/firms. Peter French (founder of JPFA) writes [s2]:

"....in 2009 we decided that the VPA was the most phonetically-grounded and detailed tool available for voice quality analysis and description..... On the basis of her (Beck's) training, it became clear that our adopting the VPA would enhance the reliability, replicability and transparency of the tests that we apply when comparing recorded voice samples..... I would estimate the number of cases in which we have used it to be in excess of 1,500. It now [i.e. during the census period] forms a significant element of the forensic speaker comparison procedure and has been critical in some cases in arriving at a positive or negative conclusion. The use of the modified VPA is mentioned in all official JPFA forensic speaker comparison reports, and it has been frequently mentioned in oral evidence delivered from the witness box in trials also."

A recent report on rigorous testing of VPA included the statement *"We further conclude that the perceptual assessment of voice quality using the VPA scheme is an essential tool in fields such as forensic phonetics"* [s3].

VPA is now a core part of the curriculum for UoY's undergraduate and postgraduate forensic phonetics programmes. VPA is also included in UoY's CPD course in Forensic Speech and Audio Analysis. Professional groups from six continents, including police, detectives, government employees, acousticians, forensic scientists in other disciplines and software engineers, have participated in the six iterations of this course, influencing forensic casework in the UK and internationally.

In Brazil a Portuguese translation of VPA ("BP-VPAS") is routinely used by Federal Police in forensic reports (~150 reports per year). Federal Police training in forensic science includes a module on VPA facilitated by PUC-SP and police from regional states confirm its value as an investigative tool [s4]. In December 2020 Brazilian government recognition of the importance of this work was demonstrated by the announcement of CAPES funding (amount currently confidential) for a strategic partnership to develop undergraduate and postgraduate programmes in forensic science which will include VPA as a core module (partner institutions = Federal Police, PUC-SP, University of Campinas).

The utilization of VPA in forensic phonetics in the UK, Europe and Brazil, drawing on our underpinning research, is supporting evidence-based decisions in criminal cases, thus contributing to the delivery of justice.

2. Commercial impact of VPA and VPA-based acoustic analysis in creative industries

As part of a planned approach to generating impact our research on acoustic classification of voices based on VPA criteria [3] has led to public and private sector investment, which has enabled the creation of a new commercial enterprise in the voice-over industry (worth \$4.4bn in 2015). Nowadays, voice-over jobs are frequently negotiated through online market places where voice-over job adverts are posted and voice-over artists audition by submitting voice samples. This creates a bottleneck for the task of shortlisting samples (often several hundred), which is a time-consuming, tedious and potentially ineffective method of voice selection. To develop a solution for this problem based on our research and associated models, we co-founded the limited company 'The Voice Distillery Ltd' in March 2020 (company number 656265). This is led by a multidisciplinary team of directors including a speech scientist (Schaeffler), a speech and language therapist (Beck) and two external directors. One is a software engineer and one an experienced voice-over artist and voice coach. The company raised private investment of £35,000 in May 2020 [s4] and has received a 'SMART Scotland award' of £67,000 in February 2020 [s5]. SMART is a Scottish business grant scheme aimed at SME's for their early stage R&D. The ongoing feasibility study develops an AI based system to automatically evaluate and rank voice auditions with respect to client brief and target population preferences, testing the reliability of such a system in reference to a VPA-based representation level. Voice Distillery (currently, in late 2020) has four employees and six freelancers associated with the company [s4]. Despite the Covid19 pandemic, the company has reached all milestones of the SMART feasibility study so far and had recorded, annotated and processed over 3000 voice samples, submitted as auditions by voiceover artists, by the end of 2020, signalling how it has begun, building on our research, to meet the needs of this sector for objective and efficient voice classification. A business development consultant reporting on the company has made specific reference to our innovative application of technology, supported by our phonetic research, to solve industry problems [s6].

3. International impact on clinical guidelines and health policy

Much of our research is designed to generate clinical impact, and our findings have directly informed international clinical guidelines for voice assessment and therapy during Covid19. Our original research [4] gave specific recommendations about which clinical acoustic voice parameters can be reliably derived from smartphone recordings. Our findings for fundamental frequency (f0), shimmer and jitter were then integrated in published clinical guidelines, written by a consortium of 11 voice and swallowing disorder experts from five countries. These guidelines [s7], published in several languages, constitute consensus recommendations from the consortium, and are meant as an international clinical guide for speech-language therapists during the Covid19 pandemic. One of the recommendations provided in the text addresses the use of mobile phones and states that "*Clinicians can use voice samples recorded through a cell phone to obtain F0*", explicitly quoting our work.

5. Sources to corroborate the impact

- [s1] [San Segundo E & Mompean J.A. \(2017\) A Simplified Vocal Profile Analysis Protocol for the Assessment of Voice Quality and Speaker Similarity. *Journal of Voice* 31\(5\):644.e11-644.e2](#)
- [s2] Testimonial from Professor Peter French, Chairman, J.P French Associates, Forensics and Acoustics Laboratory
- [s3] [San Segundo, E., Foulkes, P., French, P., Harrison, P., Hughes, V. & Kavanagh, C. \(2019\) The use of the Vocal Profile Analysis for speaker characterization: Methodological proposals. *Journal of the International Phonetic Association*. 49. 353-380. 10.1017/S0025100318000130.](#)
- [s4] Extracts from email correspondence with Professor Sandra Madureira, Pontifícia Universidade Católica de São Paulo (January 2021) in which she translates and summarises information provided by Ronaldo Rodrigues da Silva (Brazilian Federal Police) on use of VPA in forensic investigations. Personal contact details also provided.

[s5] Testimonial from Voice Distillery Ltd accountant [personal contact details]

[s6] SMART award documentation

[s7] Testimonial from Natasha Blunt, Marketing and Business Development Consultant

[s8] [Castillo-Allendes, A., Contreras-Ruston, F., Cantor, L., Codino, J., Guzman, M., Malebran, C., Manzano, C., Pavez, A., Vaiano, T., Wilder, F. and Behlau, M., Voice therapy in the context of the covid-19 pandemic; guidelines for clinical practice. *Journal of Voice*. \(pdf of early online publication, August 2020, is provided\).](#)