



HAL
open science

Seventh International Workshop on the History of Speech Communication Research

Juliusz Cęcelewski, Cécile Fougeron, Annie Rialland, Christelle Dodane, Jiayin Gao, Claire Pillot-Loiseau, Bowei Shao

► To cite this version:

Juliusz Cęcelewski, Cécile Fougeron, Annie Rialland, Christelle Dodane, Jiayin Gao, et al.. Seventh International Workshop on the History of Speech Communication Research. Seventh International Workshop on the History of Speech Communication Research (HSCR 2025), Sep 2025, Paris, France. ISCA, 2025, ISSN: 0940-6832. <10.21437/HSCR.2025>. <hal-05425726>

HAL Id: hal-05425726

<https://hal.science/hal-05425726v1>

Submitted on 19 Dec 2025

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.



HAL Authorization

Preface

The 7th edition of the Workshop on the History of Speech Communication Research (HSCR 2025) was held in Paris from September 10 to 12, 2025. This annual or biennial workshop was organized under the auspices of the ISCA Special Interest Group (SIG) on the History of Speech Communication Sciences. Since its inception in 2015 in Dresden, the workshop has been hosted in various cities, including Helsinki (2017), Vienna (2019), Prague (2021), Porto (2022), and Budapest (2024).

The 2025 edition was organized by members of the Laboratoire de Phonétique et Phonologie (CNRS & Université Sorbonne Nouvelle) and the CLESTHIA Laboratory (Université Sorbonne Nouvelle). The Laboratoire de Phonétique et Phonologie has a rich history dating back to 1911, originating from the Institut de Phonétique de l'Université de Paris. It specializes in both teaching and research in the fields of phonetics and phonology, covering a broad spectrum of areas with both theoretical and applied orientations (<https://lpp.cnrs.fr>). CLESTHIA focuses on language in interactions, discourses and society and some members on the history of speech description. The organization of this edition benefited from the support of the two organizing laboratories, the Université Sorbonne Nouvelle and its Direction des Affaires Internationales, the Association Française de la Communication Parlée (AFCP), and the project inIdEx Empirical Foundations of Linguistics (EFL).

HSCR serves as an interdisciplinary forum for researchers investigating all facets of the history of speech and communication research. This year's workshop emphasized geographical and disciplinary exchanges, underscoring the importance of collaboration within and beyond the field of speech communication.

The advancement of our field has historically been driven by intensive exchanges among scholars and students. These exchanges have been facilitated through international collaborations, the establishment of societies, the organization of congresses, and journals. Additionally, the transfer of knowledge, methodologies, technologies, and instrumentation across disciplines has played a crucial role in our progress.

The articles in this volume illustrate the richness of exchanges, both in geographical and disciplinary terms. Geographically, they span a wide range, involving Europe as well as America, Africa and Asia. The influences discussed are diverse, often bidirectional, highlighting the dialogues established across borders. Many disciplines were also involved in the birth and development of phonetics, such as medicine, physical acoustics, psycho-acoustics, linguistics, technology — notably in the realms of sound recording and film. In turn, phonetics has contributed significantly to advancements in these fields.

Nearly three years after the heart-breaking closing of the historic Institut de Phonétique de Paris building at 19 rue des Bernardins (see cover picture), which has been its home since 1928, it was a great pleasure to host this edition of the workshop on the history of our discipline in Paris. This edition was a tribute to all the people who have contributed to the Institute's outreach for so many years.

Annie Rialland and Cécile Fougeron
Paris, September 2025

TRANSFER OF KNOWLEDGE ACROSS SPACES AND DISCIPLINES: THE CASE OF CZECH PHONETICS

Pavel Šturm

*Institute of Phonetics, Charles University, Prague, Czechia
pavel.sturm@ff.cuni.cz*

Abstract: This paper explores the transfer of knowledge across geographical and disciplinary boundaries in the development of Czech phonetics during the 20th century. It focuses on how international academic mobility – including study stays abroad and internships in Prague – contributed to the formation of a dense network of intellectual exchange. These personal connections were reinforced through correspondence, the exchange of books and offprints, and participation in international congresses, which functioned as critical hubs for the circulation of phonetic knowledge. Equally important was the interdisciplinary nature of early phonetic research, which drew from linguistics, acoustical physics, medicine or scientific cinematography. In turn, phonetics enriched adjacent domains such as speech therapy, language teaching, and dialectology. By examining the case of Czech phonetics, this study highlights the broader significance of such crossings in shaping the evolution of the phonetic sciences.

1 Introduction

Phonetics has long occupied a distinctive position at the intersection of multiple disciplines, including linguistics, medicine, physics, and psychology. As a science of spoken language, its development has often relied on interdisciplinary collaboration and the adaptation of tools and methods from other fields. Equally important has been the circulation of people, ideas, and practices across national and institutional boundaries, facilitated by academic mobility, international correspondence, and participation in dedicated conferences. As Goldsmith and Laks [9] convincingly demonstrate, the evolution of scientific fields is shaped not only by specific ideas, but more fundamentally by the contexts formed by individual scholars, institutional settings, and disciplinary traditions. This paper examines Czech phonetics as a case study of how such flows of knowledge – enabled by study stays abroad, internships in Prague, and sustained intellectual exchange in a strategic location between Western and Eastern Europe – contributed to the development of both the scientific foundations of phonetics and its broader applications in areas such as speech therapy.

2 Crossing spaces

2.1 Study stays abroad

Study or research stays at foreign institutions are a highly effective means of transferring knowledge. As the following examples illustrate, such exchanges often yield mutual benefits – for both the visiting scholar and the host institution. These experiences frequently occur early in a scholar’s career, making the environment at the host institution particularly influential.

2.1.1 Josef Chlumský (1871–1939)

Chlumský began his academic training at Charles University in Prague, where he studied for two years before continuing his education abroad. His international studies included one semester in Strasbourg (then under German control), one in Berlin, and two semesters in Paris

during the academic year 1896/97. In Paris, he initially pursued Romance studies under Gaston Paris but soon developed a strong interest in phonetics, particularly through his engagement with abbé Rousselot (1846–1924), who had been teaching experimental phonetics at the Institute Catholique de Paris (ICP). During this period, Chlumský also witnessed the inauguration of the new phonetics laboratory at the Collège de France (CdF), which would soon gain international prominence. At the ICP, after completing two semesters of Rousselot's lectures and submitting an experimental phonetic thesis featuring kymographic recordings of Czech speech sounds (published as [3]), Chlumský was awarded the degree of *Maître en phonétique*.

However, a decisive moment in Chlumský's development was his second stay in Paris between 1910 and 1914. No longer a student, he collaborated closely with Rousselot, serving as his personal assistant and custodian of the laboratory at CdF. In this role, Chlumský tested new instruments, tutored Rousselot's students, conducted his own research, and contributed extensively to the *Revue de phonétique* (see [27] for an account of his numerous activities). Upon his return to Prague in 1914, he emerged as a dedicated and capable advocate of Rousselot's experimental methods, playing a key role in introducing and advancing them within the Czech phonetic tradition. Moreover, he retained a strong orientation toward French linguistic scholarship throughout his career.

2.1.2 Antonín Frinta (1884–1975)

Frinta spent three academic years studying in Prague and one year (1904–1905) in Paris, where he attended several prestigious institutions, including Université de Paris, Collège de France, and École pratique des hautes études. He studied under distinguished scholars such as Ferdinand Brunot, Antoine Meillet, Mario Roques, Jules Gilliéron, and most notably, Paul Passy (1859–1940). Under Passy's guidance, Frinta attended three key courses: a lecture on the diachronic development of French phonetics, a course on comparative phonetics of major European languages, and a seminar-style class dedicated to the analysis of phonetic texts.

Frinta adopted Passy's auditory phonetic methods and became an enthusiastic advocate of the International Phonetic Alphabet (IPA), which he applied extensively in his own work. He joined the Association in 1905 and actively contributed to its activities. Between 1906 and 1912, he published nine articles in *Le Maître Phonétique*, and in 1912 he was elected to the association's Council. Frinta developed a strong interest in pronunciation instruction (see 3.2.2). He promoted the direct method and fully embraced the association's principles in his teaching.

In 1919, Frinta was appointed as a reader at Charles University, where he began delivering phonetics lectures on a broad range of topics and languages, specializing in Slavic phonetics. His works and methods thus complemented those of Chlumský, who employed the experimental approach and focused on French and general phonetics.

2.1.3 Bohuslav Hála (1894–1970)

Hála, Chlumský's assistant in the phonetics laboratory, was sent to Strasbourg in the summer term of 1920 to improve his proficiency in French and to attend specialized courses in French phonetics alongside regular philology classes. His stay was arranged in cooperation with André Mazon, one of his lecturers during the period. This initiative indicates that Chlumský shared Frinta's belief in the importance of immersive, practical experience with a foreign language.

2.1.4 Přemysl Janota (1926–2008)

Janota belonged to a later generation of Czech phoneticians, educated after the Second World War. He was fortunate to study phonetics at Charles University before the communist regime came to power in 1948 and imposed strict limitations on Western influence. His early interest in Dutch was fostered by Bohumil Trnka, a member of the Prague Linguistic Circle. In 1947, Janota went to the University of Amsterdam, where he studied phonetics under the guidance of

Louise Kaiser (1891–1973), who had previously visited Chlumský’s laboratory during her own academic training and was personally acquainted with Hála. These existing connections made Janota’s placement in Amsterdam a natural choice. During his stay, Janota worked on his first experimental phonetic work, published in Dutch [13]. This marked the beginning of a lifelong friendship with Kaiser, as well as a sustained engagement with the Dutch language, its phonetic research traditions, and the country more broadly.

2.2 Internships in Prague

Between 1920 and 1939, as well as in the post-war period, the phonetics laboratory in Prague hosted a number of international interns and visiting scholars. They came from a range of countries – primarily France, but also Switzerland, Belgium, the Netherlands, the United Kingdom, the United States, China, Japan, Poland, Ukraine, or Yugoslavia (especially Serbia).

Most of them stayed only briefly, typically for a few days or weeks (cf. [25]). For instance, Ida Ward (1880–1949), a phonetics assistant at University College London, visited for a single day in June 1927. During her visit, her English pronunciation was recorded using a kymograph. Similarly, L. Kaiser, then an assistant at the Physiological Institute in Amsterdam, spent a week in Prague in 1923 to become familiar with the methods and research of the Prague phoneticians. During her stay, Chlumský made several kymographic recordings for her.

In contrast, some visitors stayed for several months or made repeated visits. One example is Ivan Pankevych (1887–1958), a dialectologist from Uzhhorod (present-day Ukraine), who visited the laboratory in 1922, 1923, and again in 1928. Initially, he came to learn about the available tools for dialect research, while later he used them for his own scholarly work. Among the many long-term interns, the following three individuals stand out.

2.2.1 Branko Miletić (1897–1983)

Miletić spent the academic year 1923/24 at the Prague laboratory. As one of the students of phonetics, he actively participated in all the laboratory’s classes and activities, gaining experience in physiological methods, kymographic recording, and auditory training. He attended the laboratory daily (excluding Sundays), working on his own research focused on the articulation of Serbian speech sounds – investigated through X-ray imaging – and Serbian melodic accents. This latter research formed the basis of his doctoral thesis. He remained at the laboratory for an additional year, during which he expanded his dissertation into a book [14] and completed his articulatory research for another publication [15]. Chlumský closely supported his work, reviewing and correcting Miletić’s manuscripts and research throughout his stay. Miletić returned to the laboratory several times during the 1920s and 1930s, usually for a month or two, to continue his research and make use of the laboratory’s equipment. During these visits, he also consulted on the technical requirements for establishing a phonetics laboratory in Belgrade.

2.2.2 Marguerite Durand (1904–1962)

Durand, an assistant at the Institut de Phonétique in Paris, was sent by Hubert Pernot to study under Chlumský during the summer semester of 1928/29. Her assigned project was to investigate French accentuation using experimental means. Under Chlumský’s guidance, she learned how to analyze accentuation in recordings, working with samples of her own speech, as well as those of Vendryès and two other native French speakers who happened to be in Prague at the time. She also actively participated in the broader activities of the class. The following year, Durand sent Chlumský a manuscript on French vowels, developed as a continuation of her seminar work from the previous summer. He provided detailed feedback, and the study was later published in *Revue de phonétique* [6].

2.2.3 Wu Tsung-Chi (1909–2010)

Wu, a visiting fellow from China (Chinese Academy of Science), arrived in Czechoslovakia in August 1957 and stayed until February 1958. His goal was to gain both practical and theoretical knowledge of modern experimental phonetic methods. During his stay, Wu visited 12 institutions across the country, with particular emphasis on the phonetics institute led by Hála. He met with 22 experts from various disciplines, studied general phonetics and acquired the basics of the Czech language. His training included a wide range of experimental techniques, such as spectral analysis, speech acoustics, synthesis, radiography, and other instrumental methods used in phonetic research. He also participated in fieldwork, assisting with dialect recordings. In addition to this phonetic training, he also immersed himself in the phonetic library.

2.3 Correspondence and exchange of publications

Beyond physical mobility, intellectual exchange through correspondence and the sharing of publications has long been central to academic collaboration. Letters served as a vital medium for discussing emerging ideas, critiquing methodologies and phonetic instruments, and staying informed about ongoing developments. Similarly, exchanging otherwise inaccessible journal articles and monographs between individual researchers supplemented phonetic libraries.

Chlumský corresponded with numerous prominent phoneticians and linguists, including Bogorodickij, Brunot, Chiba, Grammont, Kaiser, Mazon, Meillet, Menzerath, Millet, Panconcelli-Calzia, Pernot, Poirot, Rousselot, Scripture, and Vendryès, as well as with instrument makers such as Garten, Grandgent, Lioret, Montalbetti, and Struycken. In total, 72 international correspondents are documented (likely an underrepresentation), although the majority exchanged only a few letters. The most sustained correspondence, in terms of letter volume, was with Adrien Millet, Hubert Pernot, Georges Lote, Charles Ayer, Charles Bruneau, and E. W. Scripture. Millet and Pernot wrote over a hundred pages each, updating Chlumský on their work and on developments at several phonetic institutions in Paris (see [20] and [28]).

During the interwar period, Chlumský was not unique in this respect. Most active scholars across disciplines maintained substantial correspondence. For instance, Frinta's archive contains letters from hundreds of correspondents, including several phoneticians. However, most of the material reflects his broader interests in Slavic linguistics, literature, and culture. The collection remains largely unorganized, making comprehensive surveying difficult without browsing thousands of letters.

Frequently, books and offprints arrived along with the letters. E. W. Scripture, for instance, requested specific works from Chlumský and later asked to receive all his new publications. In return, Scripture enclosed his own writings, significantly enriching the holdings of the phonetic library in Prague, which now preserves many of his books and collected articles. This form of exchange was a common and valued means of acquiring scholarly literature.

2.4 Phonetic congresses

The International Congresses of Phonetic Sciences (ICPhS), officially established in 1932 following several smaller, locally organized events of more limited scope and disciplinary range, have since become a central part of professional life for many phoneticians, particularly those engaged in research. Czech phoneticians have been regular attendees and, incidentally, Prague holds the unique distinction of having hosted the congress twice, in 1967 and again in 2023.

The revival of the congress after World War II owes much to the persistent efforts of Hála, who advocated for its re-establishment for over a decade. It was partly thanks to him that the gap between the last pre-war congress in 1938 – which he personally attended – and the first post-war event in 1961 was ultimately bridged (see [26]). Hála understood well the value of such international gatherings, especially at a time when the communist regime severely restricted contact with researchers abroad. It is important to emphasize that the congresses were

more than venues for scientific exchange through lectures and discussions; above all, they provided opportunities to connect with colleagues and form new relationships, many of which were sustained through ongoing correspondence (see above).

3 Crossing disciplines

The previous sections focused on the transfer of (phonetic) knowledge across individuals, laboratories, and national boundaries, whether by means of physical relocation or remote correspondence. Section 3 will examine cases in which knowledge is transferred between disciplines or scientific environments. In this context, phonetics may be seen both as a field enriched by the insights and methodologies of other disciplines and as one that contributes its own knowledge and practices to them. Again, examples will be taken from Czech phonetics.

3.1 Enrichment of phonetics

3.1.1 Tuning forks and acoustic physics

Chlumský was captivated by Rudolph Koenig's Grand Tonometre, an elaborate and highly precise set of tuning forks acquired by Rousselot following Koenig's death. Chlumský set out to build a precision replica of the instrument which, while smaller in scope, featured 100 tuning forks covering a range from 16 Hz to 3,296 Hz (see [23] for further details).

Neither Rousselot nor Chlumský possessed the full range of skills required to construct a tonometer. The project demanded a specialized and interdisciplinary body of knowledge, combining artisanal craftsmanship, metallurgy and acoustic physics. Steel had to be selected with the same care as wood for musical instruments, and each fork had to be cast in a single piece, then painstakingly shaped and tuned for precision. Crucially, the project depended on a network of experts: R. Montalbetti played a key role in the production process, while a mechanic from Pirard et Coeurdevache was responsible for the finishing work. In 1921, Chlumský travelled to Paris to supervise this intricate process. He often collaborated with Rousselot and Millet, comparing the newly crafted forks with Koenig's originals to ensure accuracy. The calibration and testing of the forks took several months of daily work for just thirteen forks.

Tuning forks had practical applications in experimental phonetics. Chlumský employed them mostly as a pedagogical tool, training phonetic students and dialectologists to perceive subtle vowel distinctions. In the 1930s, Hála further utilized them for an extensive acoustic analysis of Czech vowels [10].

3.1.2 Medicine

The development of medicine in the 19th century was a key factor to the emergence of phonetics at the end of the century, when integrated with acoustic physics and, most importantly, linguistics [1]. Anatomical and physiological observations by 19th-century physicians and scholars significantly enriched phonetic descriptions of articulatory processes, grounding them in empirical evidence. Techniques such as vivisection, physiological experimentation, and measurement provided a new level of certainty to earlier speculative accounts. Instruments originally developed for physiological research were increasingly repurposed for the study of speech.

Phonetics was notably advanced by the contributions of two Czech physiologists: Jan Evangelista Purkyně (1787–1869) and Jan Nepomuk Czermak (1828–1873) (see [24]). Purkyně published a remarkably systematic physiological account of the articulation of speech sounds [19] two decades before Brücke's *Grundzüge der Physiologie* appeared [2]. Czermak, for his part, improved the laryngeal mirror, enabling direct observation of the vocal folds and effectively founding the field of clinical laryngoscopy. However, his work went beyond medical diagnostics: he also applied this instrument to the study of speech sounds produced at the glottis (e.g. [4]) and delivered popular lectures on the topic of "voice and speech" [5].

Collaboration between phonetics and medicine remained strong in the Czech lands. Emanuel Kovář (1861–1898), a linguist and self-taught phonetician, studied philology at the Faculty of Arts while also attending selected courses at the Faculty of Medicine. As a lecturer, he taught not only linguistics but also speech disorders to philology students. Deeply interested in the relationship between brain function and speech, Kovář was particularly drawn to emerging research on neurological impairments. His promising career was unfortunately cut short by suicide.

Later, Hála continued to cultivate strong ties to the medical community. He collaborated with Miloš Sovák (1905–1989), an otorhinolaryngologist specialized in phoniatrics who played a central role in the development of modern Czech speech therapy (see 3.2.1). Together, they authored a popular book on voice, speech, and hearing, which went through several editions [12]. Hála also maintained professional contacts with other notable figures in the field, including Raoul Husson (1901–1967).

3.1.3 Scientific cinematography

The desire to observe the activity of the vocal folds during phonation dates back centuries (cf. [16]). While Jan Nepomuk Czermak's improvements to the laryngeal mirror made direct visual examination of the vocal folds possible, the next major step involved capturing their rapid, cyclical movements on film. A significant advancement was achieved by Bohuslav Hála in 1928–1929 through his collaboration with Ludovít Honty, a specialist in scientific cinematography. Their work marked a breakthrough in visualizing glottal activity, combining high-speed cinematography with stroboscopic techniques, as detailed in [20]. The resulting footage, featuring Hála's own vocal folds, was presented to an international audience in 1930 during the Fourth International Congress of Logopaedics and Phoniatrics, held in Prague. The film was widely acclaimed, and a French-language report by Hála and Honty from 1931 remains a frequently cited source in the literature on phonation [11]. Panconcelli-Calzia recognized it as the first successful instance of capturing the vocal folds using high-speed cinematography [18].

3.1.4 Sound recording

One of the most essential technical advancements for phonetic research was the ability to record sound onto a physical medium and play it back for repeated analysis (auditory or instrumental). Early attempts, such as the use of the phonograph with wax cylinders, laid the groundwork, but it was the advent of more durable and reliable gramophone disc recordings that enabled broader distribution and systematic archiving. During the early 20th century, several speech archives were established across Europe, in cities such as Vienna, Paris, and Berlin. As early as 1907, Frinta participated in a phonographic training course organized by the University of Vienna and the Academy of Sciences, where he learned to record speech on specialized gramophone records for dialectological purposes, although he never really used that skill. It was not until the late 1920s that systematic efforts to archive Czech speech were undertaken.

Chlumský collaborated with French phoneticians and the Pathé recording company to initiate the first major wave of speech recordings in Czechoslovakia in 1929. The recorded material included dialect samples, public speeches, poetic recitations, or music. The endeavour was logistically demanding and required unprecedented sums of money. It was supported by the Czechoslovak Academy of Sciences and Arts and its Phonographic Commission, chaired by Chlumský. On the French side, the project was strongly supported by Hubert Pernot, a key collaborator. Together, they worked to overcome a range of technical challenges, including issues with the reliability of wax discs, which affected the consistency and quality of recordings.

3.2 Benefit of phonetics

This section turns to disciplines that have drawn on the findings of phonetics or adopted its methods to investigate spoken language. These fields often concentrate on specific contexts or populations – such as atypical speech, non-native speech, or dialectal variation – where the

application of phonetic analysis proves especially valuable. The final section addresses phonology, a discipline formally distinguished from phonetics during the interwar period, yet one that remains closely interconnected with it in both theory and practice.

3.2.1 Speech and language therapy (SLT)

Both Rousselot and Chlumský had significant ties to the treatment of atypical speech and hearing. Rousselot for instance diagnosed hearing loss by tonal testing with tuning forks, evaluating both the range and severity of impairment. In Prague, this connection was evident in the education and training of professionals in the field. Medical practitioners frequently visited the phonetic laboratory, where they learned to apply experimental phonetic research in clinical settings and address various speech disorders, including oesophageal speech.

A key figure in the development of Czech SLT was Hála, especially because of his contributions to the education and professional training of therapists. His influence extended partly from his phonetic lectures and published works. Hála's physiological studies on the voice had a lasting impact, informing later monographs on voice pathology [22]. His acoustic analyses of vowel characteristics, such as volume and carrying power, played a key role in auditory testing methods developed by his student, Přemysl Janota. Another prominent student, Miloš Sovák, became a key figure in SLT and advanced the field considerably.

Hála's primary influence, however, lay in his extensive extracurricular teaching. He organized numerous training sessions, courses, and workshops that combined practical applications with research-based insights. Hála delivered special lectures and demonstrations for teachers of the deaf and provided phonetics instruction for staff at the SLT Institute in Prague. He also participated in SLT training programs for both schoolteachers and preschool teachers, emphasizing the importance of early speech development and preventive phonetic care. One of his students, Karel Ohnesorg, went on to publish a foundational longitudinal study on children's speech development [17]. Hála's efforts were not confined to Prague, as he also travelled widely across the country, bringing phonetic education to rural areas and smaller institutions.

3.2.2 Teaching of foreign languages

The intersection of phonetics and language teaching was most clearly embodied in the work of Frinta. A dedicated secondary-school teacher of French, he retained this role even after his appointment as a lecturer at Charles University in 1919. He was also an active contributor to a local Czech journal *Lectures françaises*, where he published accounts of French pronunciation and examples of phonetic transcription under the column "The Phonetic Corner". Frinta was among the first educators to incorporate the gramophone into language instruction, using it to provide students with authentic pronunciation models (especially in terms of the prosody of speech). Chlumský likewise underscored the significance of this device for language teaching, even offering university lectures dedicated to the topic.

In 1909, Frinta published a monograph on the teaching of foreign pronunciation [7], aiming to improve prevailing practices at Czech schools. He criticized the limited attention given to pronunciation instruction and the widespread tolerance of strong Czech accents in foreign language classes. He held both students and teachers accountable for these shortcomings: students, for lacking motivation, and teachers, for failing to provide accurate pronunciation models and lacking a solid grounding in phonetics. Frinta emphasized the need for teachers to possess both practical and theoretical phonetic knowledge and recommended summer phonetics courses, which also allowed for language immersion in natural contexts. Most importantly, he advocated for the consistent use of phonetic transcription (IPA) as a central tool in pronunciation teaching. He argued that its benefits should be clearly explained to students, and that traditional orthography should be excluded from pronunciation instruction altogether.

3.2.3 Dialectology and language change

Frinta's work was instrumental in establishing the theoretical and methodological underpinnings of linguistic geography in the Czech context. He drew inspiration from scholars such as Jules Gilliéron and appreciated the meticulous, fieldwork-based approach of Rousselot [21]. Frinta perceived dialects as regionally varied forms of a national language, noting that while the boundaries between dialects were often fluid, it was possible to map precisely the distribution of individual pronunciation features (using *isophones*).

Phonetics played a central role in Frinta's dialectological research, as it enabled the direct observation and documentation of spoken language in real-world settings. He insisted that first-hand experience in local environments was essential, as questionnaires alone could not capture the nuance and variability of regional speech. His own fieldwork took him across various Slavic countries, including the Czech lands, Poland, and the Balkans, where he gathered empirical material that informed his studies of pronunciation and sound change. By focusing on phonetic phenomena such as the articulation of /v/, he was able to trace patterns of linguistic variation and link them to historical processes.

One of his most significant methodological contributions was the use of isophones to map the "colourful diversity" of spoken language. Layering numerous isophones onto a single map revealed complex patterns of variation that would have remained invisible in a purely theoretical approach. This method also allowed him to make well-supported reconstructions of older language stages. For instance, in his habilitation thesis [8], Frinta analyzed the current phonetic forms of labial consonants across Slavic regions to hypothesize the probable articulation of the Proto-Slavic /v/, which he concluded was bilabial. His insistence on beginning historical reconstruction from the detailed phonetic reality of contemporary speech exemplified his belief that dialectology must be grounded in the living language, not in written or standardised forms.

3.2.4 Phonology

In the interwar period, Prague was home to two prominent approaches to the study of spoken language. One was the *phonetic school* led by Chlumský, centred around his Laboratory of Experimental Phonetics, where Hála served as his assistant. Frinta, while deeply engaged in auditory phonetics, did not establish a formal school of his own but belonged to a loosely connected group of phoneticians and linguists working independently. In contrast, the second major body was the Prague Linguistic Circle (PLC), which developed a more comprehensive theoretical framework known as *functional structuralism*. Within this framework, the functional perspective was emphasized by Vilém Mathesius (1882–1945), the Circle's founder and leader, while the structural component was championed by N. S. Trubetzkoy (1890–1938) and Roman Jakobson (1896–1982), the principal architects of phonology.

From the outset, Trubetzkoy strongly insisted on a strict conceptual separation between phonetics and phonology, a stance articulated clearly in his early writings (e.g. [29]). This division, at times marked by rivalry or even tension, may have been necessary, as phonology was emerging as a distinct scientific discipline. Officially recognized as such at the First International Congress of Linguists in The Hague in 1928, phonology had to assert its legitimacy and differentiate itself from the already well-established field of experimental phonetics. The institutional imbalance at the time reflected this struggle: while there were two university chairs in phonetics (held by Chlumský and Frinta), there was no corresponding academic position for phonology.

Trubetzkoy's attitude toward phonetics was more nuanced than is sometimes assumed. While he famously argued for a strict separation between phonetics and phonology (going so far as to characterize phonetics as a biological science), his scepticism was primarily directed at the experimental phonetics practiced by figures like Panconcelli-Calzia. This type of phonetics, focused on instrumental and physiological measurement, struck Trubetzkoy as excessively empirical and disconnected from linguistic function. However, he did not reject phonetics as a

whole. On the contrary, he acknowledged the importance of auditory phonetics and phonetic transcription as essential steps in phonological analysis. In correspondence with Jakobson, Trubetzkoy even praised the British tradition for practical phonetic descriptions of exotic languages, noting that although it lacked theoretical sophistication, its findings could be easily translated into the language of phonology. His critique, then, was not of phonetics per se, but of a particular version of it that he saw as incompatible with structural analysis.

Mathesius, on the other hand, maintained a consistently positive relationship with phonetics throughout his career. A distinguished scholar of English, Mathesius was well-read in both experimental and auditory phonetics. He was particularly influenced by the practical orientation of scholars such as Sweet, Jespersen, Viëtor, and Passy, but he also intimately knew the experimental work of Rousselot, Meyer or Scripture. Mathesius was a member of the IPA, and he held the work of the Frinta in high esteem. Though primarily a linguist and a phonologist, Mathesius clearly saw phonetic analysis as an indispensable tool for understanding language structure.

Chlumský held a markedly different view on the relationship between phonetics and phonology. He regarded phonology not as an independent discipline but as a subfield of phonetics – specifically, a third dimension complementing the articulatory and acoustic aspects of speech sounds. From this perspective, he argued that phonology offered nothing fundamentally new, and that phonetics already encompassed a more complete and comprehensive study of speech. However, Chlumský misunderstood the aims and scope of phonology, interpreting it primarily as the psychology of speech sounds rather than as the analysis of sounds within the structure of a language system. As a result, central phonological concepts such as the phoneme, functional opposition, and the notion of phoneme realization were largely foreign to his thinking.

4 Conclusions

Czech phonetics was used as a case study to illustrate how knowledge circulated within the early phonetic community in Europe. A key factor in the development of individual phoneticians was their early exposure to phonetics in international settings, where specific institutions, people, and contexts shaped their future careers. At the same time, the Prague laboratory served as an important hub for training foreign students, particularly from Central and Eastern Europe, offering hands-on experience in experimental phonetics. Personal correspondence between scholars often sustained further long-term intellectual exchange and may have been more influential than the relatively infrequent phonetic conferences.

Another central theme was the inherently interdisciplinary nature of phonetics. The field has long drawn on tools and concepts developed in other domains – such as tuning forks from acoustical physics or kymographs from medical research – and repurposed them for the study of speech. However, phonetics was also a source of inspiration for other disciplines that worked with human speech. This was most evident in dialectological studies and phonology, which needed precise phonetic descriptions, or in speech therapy and foreign language instruction, where phonetic knowledge and tools were crucial for successful acquisition.

5 Acknowledgments

This publication was supported by the Cooperatio program of Charles University, Linguistics research area, implemented at the Faculty of Arts of Charles University.

6 References

- [1] ASHBY M. & J. MCELVENNY. 2022. The emergence of phonetics as a field. In *Interviews in the History of Linguistics*, 51–60. Language Science Press.
- [2] BRÜCKE, E. 1856. *Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Taubstummlehrer*. Carl Gerold's Sohn.

- [3] CHLUMSKÝ, J. 1902. Analyse du courant d'air phonateur en tchèque. *La Parole* 4. 130–143, 395–406, 483–496, 665–683, 736–741.
- [4] CZERMAK, J. N. 1879. Über den Spiritus asper und lenis, und über die Flüsterstimme, nebst Bemerkungen zur phonetischen Transscription der Kehlkopflaute. In *Gesammelte Schriften I*, 750–766. Wilhelm Engelmann.
- [5] CZERMAK, J. N. 1879. Stimme und Sprache. In *Gesammelte Schriften II*, 60–104. Wilhelm Engelmann.
- [6] DURAND, M. 1929. Étude sur les phonèmes postérieurs dans une articulation parisienne. *Revue de phonétique* 6. 241–256.
- [7] FRINTA, A. 1909. O učení cizí výslovnosti [On teaching foreign pronunciation]. In *Pátá výroční zpráva reálky v Mladé Boleslavi*. Hejda a Zbroj.
- [8] FRINTA, A. 1916. *Fonetická povaha a historický vývoj souhlásky „v“ ve slovanštině [Phonetic Nature and Historical Development of “v” in Slavic]*. Česká akademie.
- [9] GOLDSMITH, J. & B. LAKS. 2019. *Battle in the Mind Fields*. University of Chicago Press.
- [10] HÁLA, B. 1941. *Akustická podstata samohlásek [Acoustical Basis of Vowels]*. ČAVU.
- [11] HÁLA, B. & L. HONTY. 1931. La cinématographie des cordes vocales à l'aide du stroboscope et de la grande vitesse. *Otolaryngologia Slavica* 3. 1–12.
- [12] HÁLA, B. & M. SOVÁK. 1941. *Hlas, řeč, sluch [Voice, Speech, Hearing]*. Česká grafická unie.
- [13] JANOTA, P. 1947. *Onderzoek betreffende de kwantiteit van de Tsjechische en Nederlandse klinkers [Research on the Quantity of Czech and Dutch Vowels]*. University of Amsterdam.
- [14] MILETIĆ, B. 1926. *O srbo-chorvatských intonacích v nářečí štokavském [On Serbo-Croatian Intonations in the Štokavian Dialect]*. FF UK.
- [15] MILETIĆ, B. 1933. *Izgovor srpskohrvatskih glasova [Pronunciation of Serbo-Croatian Sounds]*. Srpska kraljevska akademija.
- [16] MOORE, P. 1991. A short history of laryngeal investigation. *Journal of Voice* 5(3). 266–281.
- [17] OHNESORG, K. 1948. *Fonetická studie o dětské řeči [A Phonetic Study on Child's Speech]*. FF UK.
- [18] PANCONCELLI-CALZIA, G. 1931. [Part of Bibliographia Phonetica]. *VOX* 17(1). 77–78.
- [19] PURKYNĚ, J. E. 1835 [1958]. *Pojednání přírodoslovné o mluvě lidské [A Natural Science Treatise on the Human Speech]*. In *Jan Evangelista Purkyně – Opera Omnia (Volume XII)*, 89–113. ČSAV.
- [20] RIALLAND, A., P. ŠTURM & C. FOUGERON. 2024. Rousselot's succession through Pernot and Millet's correspondence to Chlumský. In *Proc. Sixth International Workshop on the History of Speech Communication Research*, Budapest, 25–34. TUDpress.
- [21] ROUSSELOT, P.-J. 1891. *Les modifications phonétiques du langage étudiées dans le patois d'une famille de Cellesrouin (Charente)*. H. Welter.
- [22] SOVÁK, M. 1945. *Stroboskopický výzkum hlasové pathologie [Stroboscopic Research in Voice Pathology]*. ČAVU.
- [23] ŠTURM, P. 2015. The Prague historical collection of tuning forks: A surviving replica of the Koenig tonometre. In *Proc. First International Workshop on the History of Speech Communication Research*, Dresden, 95–105. TUDpress.
- [24] ŠTURM, P. 2019. The contribution of Czech phonetics to laryngeal investigation. In *Proc. 19th International Congress of Phonetic Sciences*, Melbourne, 1903–1907. ASSTA.
- [25] ŠTURM, P. 2022. The Laboratory of Experimental Phonetics in Prague between 1919 and 1939: Research practices and student instruction. In *Proc. Fifth International Workshop on the History of Speech Communication Research*, Porto, 75–85. TUDpress.
- [26] ŠTURM, P. 2023. The 6th ICPHS in Prague and the organized life of phonetics. In *Proc. 20th International Congress of Phonetic Sciences*, Prague, 3231–3235. Guarant International.
- [27] ŠTURM, P. 2025. From Paris to Prague: Joseph Chlumský's enduring commitment to Rousselot's experimental phonetics. In *Actes du Colloque en hommage à l'Abbé Rousselot*. Société de linguistique de Paris.
- [28] ŠTURM, P., A. RIALLAND & C. FOUGERON. 2024. Pernot, Scripture and Millet: Different ideas for running a journal of experimental phonetics in the early 20th century. In *Proc. Sixth International Workshop on the History of Speech Communication Research*, Budapest, 35–45. TUDpress.
- [29] TRUBETZKOY, N. S. 1929. Zur allgemeinen Theorie der phonologischen Vokalsysteme. *Travaux du Cercle Linguistique de Prague* 1. 39–67.

EARLY TRANSCRIPTIONS OF ENGLISH AND FRENCH IN THE INTERNATIONAL PHONETIC ASSOCIATION: THE WEIGHT OF SPELLING REFORM ON THE BIRTH OF THE IPA

Jacques Durand¹, Chantal Lyche²

¹CLLE, Université Toulouse-Jean Jaurès, France, ²University of Oslo, Norway
jacques.durand@univ-tlse2.fr

Abstract: The establishment of the International Phonetic Association and the creation of its alphabet (the IPA) attest to the fact that scientific advancement emerges from collective collaboration and sustained debate. The development of the IPA is closely intertwined with progress in the fields of phonology and phonetics. However, the founders of the Association—such as Paul Passy, Otto Jespersen, and Wilhelm Viëtor—alongside influential scholars like Isaac Pitman, Alexander Ellis, and Henry Sweet, were also ardent advocates of spelling reform. We examine the extent to which this reformist agenda influenced the evolution of the IPA, with particular focus on debates within the *Maître Phonétique* concerning the concepts of ‘phonetic transcription’ and ‘phonetic spelling.’ We show that spelling reform was consistently viewed as a core concern and constituted a persistent, underlying thread in the evolution of the IPA. We argue that much of the debate between Sweet, Passy and Jones in *Le Maître Phonétique* had no winner as such. However, the adoption of Sweet’s ‘broad’ transcriptions preferentially based on Latin-based symbols identically used across languages led to a confusion between the task of phonology/phonetics, the devising of spelling systems and the pedagogical applications of the IPA.

1 Introduction

The establishment of the International Phonetic Association and the creation of its alphabet (the IPA) attest to the fact that scientific advancement emerges from collective collaboration and sustained debate. The development of the IPA is closely intertwined with progress in the fields of phonology and phonetics. However, the founders of the Association—such as Paul Passy, Otto Jespersen, and Wilhelm Viëtor—alongside influential scholars like Isaac Pitman, Alexander Ellis, and Henry Sweet, were also ardent advocates of spelling reform. After a brief examination of this historical background, we study the foundation of the Association and show that there was much ambiguity between the devising of an international phonetic alphabet for scientific purposes, spelling reform and pedagogical applications. The controversies among various figures—particularly between Sweet and Passy/Jones within *Le Maître Phonétique*—had a profound impact on the development of the IPA. The push for a quasi-universal use of Latin-based symbols to represent similar, yet often distinct, sounds across languages still echoes in contemporary practice. Moreover, the adoption of ‘broad’ transcriptions featuring word blanks and standard punctuation has long contributed to a blurring of the line between phonological representations and orthographic systems.

2 Historical background

The development of the International Phonetic Alphabet (henceforth IPA) was a collaborative endeavour, it was “hundreds of people around the globe reading and discussing every month topics in phonetics and language-teaching.” [1] Nonetheless, it was Paul Passy (1859–1940)

who played a pivotal role in steering the project forward. His unwavering commitment culminated in the establishment of an organization that would later become the *Association Phonétique Internationale* in 1897 and, subsequently, the International Phonetic Association in 1970.

The concept of a universally applicable phonetic alphabet was by no means novel in the late 19th century. Historical evidence indicates that the idea of one-to-one correspondences between phonemes and graphemes has ancient origins, frequently underlying early alphabetic writing systems [2], [3]. Universal alphabets characterizing speech sounds in terms of phonetic features were already being devised in the 17th and 18th centuries [4]. Notably, in 1686, Francis Lodwick (1619–1694), a Flemish scholar, published *An Essay toward a Universal Alphabet* [5]. In this work, he articulated a visionary rationale for such an alphabet. Lodwick proposed that a universal phonetic system could facilitate children's acquisition of foreign pronunciations and enable the accurate representation of any language's pronunciation. He wrote:

It will also be useful to perpetuate the true Sounds of any Language, and serve as a Standard thereof to after Ages: for if all the single Sounds expressable, be here Collected and Characterized, *And that no one Character have more than one Sound, nor any one Sound be expressed by more than one Character*; it cannot fall out that any Character should be falsly pronounced, but it will soon be discovered; for this false Sound he giveth it, must be the true Sound of some other Letter of this Alphabet: and so none can Erre herein, but he that wilfully or carelesly will do it. [5, p. 3, original spelling]

In 18th-century France, two additional projects merit mention. Charles de Brosses (1709–1777) devised an “organic universal alphabet” primarily intended for cross-linguistic comparisons [6], while Jean-Baptiste Montmignon (1737–1824) developed a “figurative alphabet applicable to all languages” [7], echoing many of Lodwick’s objectives [2], [3].

Nineteenth-century initiatives to create a universal, phonetically based alphabet thus drew upon this long-standing intellectual tradition concerned with the systematic transcription of human language. Among the newly developed systems, those that remained closest to familiar European alphabets tended to be the most widely accepted. Two prominent examples illustrate this trend. François Volney advanced several proposals and, in 1804, published the influential *L'alphabet européen appliqué aux langues latines*. He also left a bequest of 24,000 francs to the *Institut de France*, instructing that it be awarded to any individual capable of devising a “harmonic alphabet” based on Latin script—one that would bring coherence to the fragmented orthographic practices of the time. Another significant contribution came from the German Egyptologist Carl Richard Lepsius, who published his *Standard Alphabet* in 1854 (German edition), followed by an English edition in 1855. Intended as a uniform orthographic system based on ‘European’ characters, the alphabet was designed for use with both unwritten languages and transcriptions of non-European scripts. Despite its scientific rigor and success—particularly in Africa—the system’s reliance on an extensive set of diacritics discouraged widespread adoption among users and publishers.

Thus, while the idea of a universal phonetic alphabet was very much in keeping with the intellectual spirit of the 19th century, its broader success depended not only on conceptual soundness, but also on the presence of influential advocates capable of demonstrating its feasibility to a larger scholarly and pedagogical community. Paul Passy would take up this challenge with unwavering dedication, joined by international colleagues who shared his conviction in the necessity of such an alphabet.

3 The birth of the IPA

Let us briefly recall that Passy, as early as 1882-1883, had published *L'Anglais parlé, ou Méthode pratique pour apprendre à parler, à comprendre et à lire en anglais sans maître* (154 pages in two volumes) under the imprint of Eizak Pitman, Fonetic Institute [sic]. The originality of this publication lies not so much in the contents of the lessons but in the complete adoption of a phonetic transcription based on Pitman's 'phonotypic alphabet' for all the English examples instead of the standard spelling. This choice marks the first entry by Passy into the world of modern phonetics. Passy, determined to improve foreign language teaching, strongly believed that learners would quickly acquire an adequate pronunciation if presented with texts in phonetic spelling rather than in standard orthography. Passy created on 2 January 1886 the *Association phonétique des professeurs d'anglais*, soon after it was enlarged to include all language teachers and became *Dhi Fonètik Tîtcierz' Asóciécon* with its journal *Dhi Fonètik Tîtcer* first published in May 1886. The International Phonetic Alphabet (IPA) then began modestly—compared to Lodwick's universal alphabet—, with a relatively limited initial objective: the transcription of English. Its broader development was catalysed by Paul Passy's participation in the Stockholm philological meeting in August 1886, during which the *Quousque Tandem Scandinavian Association for the Reform of Language Teaching* was established. The *Association's* first proposal emphasized the primacy of oral language in language instruction, advocating for the early introduction of phonetic transcriptions of texts rather than reliance on conventional orthography. Passy, along with other proponents like Otto Jespersen, sought to extend this approach to all languages. However, in order to build broader consensus, the initial implementation was limited to languages characterized by a significant divergence between standard spelling and actual pronunciation. The connection between the *Association* and language teaching remained robust for a long period. In January 1889, *The Phonetic Teacher* was renamed *Le Maître Phonétique*, and at the same time, *The Phonetic Teacher's Association* became *L'Association Phonétique des Professeurs de Langues Vivantes*, with Wilhelm Viëtor serving as its first president [8]. In 1897, the *Association* expanded its scope and adopted the name *Association Phonétique Internationale*, reflecting its increasingly international membership, which by then included 865 members from a diverse range of countries. The most recent structural revision occurred in 1970, when the organization formally became the *International Phonetic Association*, and *Le Maître Phonétique* was succeeded by the *Journal of the International Phonetic Association*.

Paul Passy consistently placed language instruction at the center of his vision for phonetics. In a 1929 analysis of the main applications of phonetic science, he emphasized the importance of foreign language instruction concluding: "One could nearly say that without Phonetics, modern languages seem dead; with Phonetics, dead languages seem alive" [9: 9]. His international collaborators (such as Jespersen, Viëtor, and Lundell, to name a few) were all prominent phoneticians and language educators who largely viewed the phonetic alphabet as a pedagogical tool. Jespersen, for instance, had suggested to Passy the renaming of the organization to *L'Association Phonétique des Professeurs de Langues Vivantes* [10: 12]: "Would it not be preferable to have an Association of teachers of *all modern* languages? The same considerations that plead in favour of the phonetic method for the teaching of English, indeed make it a necessity, apply equally well to the teaching of French, Russian, Danish, etc., etc."

However, language pedagogy was not the sole impetus behind Passy's foundational work that eventually led to the development of the International Phonetic Alphabet. Passy's ambition extended beyond this practical application and also envisaged the phonetic alphabet as a possible alternative to traditional orthography although the first issue of *The Phonetic*

Teacher (May 1886) asserts the opposite¹: “There are lots of spelling reform papers already, what’s the use of having one more? True, reader; but this isn’t going to be “one more”. We have nothing to do with spelling reform; individually, we may favour it, but as a body, we know nothing about it. Listen and I’ll tell you who and what we are, and why we start a paper.” Despite this disclaimer, the fact that both *The Phonetic Teacher* and *Le Maître Phonétique* (henceforth MP) were published entirely in phonetic transcription underscores this broader objective. It is therefore noteworthy that, although the IPA was regarded as useful for linguistic description, this was always a secondary consideration for Passy. In contrast, the eminent phonetician Pierre-Jean Rousset (1846–1924)—who notably refused to adopt the IPA [8]—was motivated solely by linguistic concerns when designing his own alphabet. However, with its 288 possible vowels, his system proved too complex and was soon forgotten [11].

4 The spelling issue

Passy's experience in language teaching led him to be the driving force behind two closely related endeavours: a spelling reform and the development of a phonetic alphabet. In doing so, he joined a broader movement of the time that advocated for the reform of standard spelling.

His work was significantly influenced by British pioneers, particularly Isaac Pitman (1813–1897) and Henry Sweet (1845–1912). Pitman, himself an educator, recognized the detrimental effects of conventional English orthography and began reimagining English transcription based on its phonological structure. This initiative resulted in the creation of a phonetic shorthand system, published in 1837 under the title *Stenographic Sound-Hand*. Subsequently, Pitman collaborated with Alexander Ellis (1814–1890) to develop a “phonotypic alphabet”, which proved to be a major influence on Passy. In particular, this alphabet introduced several symbols outside the traditional Latin set—some of which would later be incorporated into the IPA with their modern phonetic values. These include [ʃ] and [ʒ] for the sounds in *vicious* and *vision*, respectively, as well as [ŋ] for the <ng> in *sing*. Other symbols—such as precursors to the modern IPA [ɛ], [ɣ], and [ʊ]—were also present, although they initially had different phonetic values. A variant of this alphabet was devised by Passy and appeared in the first issue of *The Phonetic Teacher* (May 1886), which opened with the line: “Dher are lots of spelin reform pépers ôlredi; hwot’s di yûs ov havin wœn mór?”

Sweet followed in Pitman’s footsteps in his pursuit of a standard orthography more closely aligned with actual pronunciation. His first major work, *A Handbook of Phonetics*, published in 1877, carried the subtitle *A Popular Exposition of the Principles of Spelling Reform*, reflecting this commitment. The *Handbook* had a profound influence on Passy and played a foundational role in shaping the principles underlying the IPA. Sweet, who joined the *Association Phonétique* in 1886 and was appointed as Honorary President of the *Phonetic Teacher’s Association* in January 1888, fully supported the younger generation’s initiative to introduce phonetic texts in the classroom before standard orthography. In his *Handbook*, he also advocates for the need for spelling reforms: “Of course the orthographies of most of the Continental languages require reform as well as English; French, especially, most urgently demands a thorough change. Indeed, there is no reason why foreigners should not learn French on a phonetic system, leaving the present French spelling to be acquired afterwards, even if the French themselves do not inaugurate a reform.” [12: 198] In his exploration of several transcription systems (organic, Romic, glossic) he favours the Romic (or Romanic), Latin-based, which the IPA is partially based on.

¹ The quote is given here in standard orthography.

Sweet defended the necessity of separating two types of transcription, a ‘broad’ and a ‘narrow’ one, which anticipated the modern concept of phoneme: “Again, among the vowels of any one language we must distinguish between those differences which are *distinctive* [underlined in the original], that is to which differences of meaning correspond, and those which are not. Thus, the first element of the diphthongs in ‘by’ and ‘out’ vary considerably: some people sound them broad as in ‘father,’ some thin, as in ‘man,’ with various intermediate sounds. And yet the meaning of the words remains unchanged. The distinction between the vowels of ‘men’ and ‘man,’ on the other hand, though really slighter than that of the different pronunciations of ‘by’ and ‘out,’ is a distinctive one” [12: 182]. This position is strongly supported by Passy and systematically applied in the MP and in his analyses of French [8]. However, Sweet in his objective to “make phonetic spelling less repellent to the general public” [13: 73] made proposals that Passy and Jones strongly refuted.

5 Debates on notation

The elements sketched above provide the broad background against which debates on notation took place within the *Maître Phonétique*. Passy as general editor was instrumental in determining which direction the debate would take. From more or less the beginning, the line he followed was to advocate the same phonetic notation both as the medium of communication between members and as a pedagogical tool providing reading material in the classroom. The conflation of these two functions explains ambiguities that were never solved. For instance, whereas the word was considered not to have any phonetic reality, transcribed utterances preserved blank spaces between words going even beyond traditional divisions as we shall see.

To demonstrate our point, let us take an example from the ‘Partie des élèves’ (Learners’ Corner) from 1894 [14] which uses the transcription adopted by Passy in all his contributions in English to this journal, preceded by our transliteration. This passage provides an illustration of the kind of material often included by Passy for beginners. We will not focus on the advantages and possible drawbacks of such texts but note that, as time went on, fewer and fewer pedagogical samples were submitted to the MP, a problem regularly bemoaned by Passy.

The woman and her pig

An old woman was sweeping her house one morning, and she found a little crooked sixpence. “What shall I do”, said she, with this little sixpence? I’ll go to market, and buy a little pig, and come home tonight to cook my apple-dumplings.”

So she went to market, and bought a little piggy-wiggy. As she was coming home, she came to a stile; the pig would’nt go over the stile. She said “pig, pig, go over the stile, or I shan’t get home tonight, to cook my apple-dumplings, oh.” – but the pig wud’nt.

ðə wʊmən ənd ər piɡ

ən ould wʊmən wəz swi:pɪŋ ər haus wʌn mɔ:rnɪŋ, ən ʃi faʊnd ə litl krukɪd sɪkspəns. “hwɒt ʃl ai duw”, sed ʃi, “wið ðis litl sɪkspəns? ai l gou tə mɑ:rkɪt, ən bai ə litl piɡ, ən kʌm hoʊm tənaɪt tə kuk mai æpl-dʌmplɪŋ.”

sou ʃi went tə mɑ:rkɪt, ən bɔ:t ə litl piɡi-wiɡi. æz ʃi wəz kʌmiŋ hoʊm, ʃi keɪm tə ə stɑɪl; ðə piɡ wud nt gou ɔʊvəɪ ðə stɑɪl. ʃi sed “piɡ, piɡ, gou ɔʊvəɪ ðə stɑɪl, ər ai ʃənt get hoʊm tənaɪt, tə kuk mai æpl-dʌmplɪŋz, ɔʊ.” – bət θə piɡ wud nt.

From a phonetic point of view, a first observation is that by 1894 the symbols were more or less identical with the ones used today. We can immediately see that this transcription of English adopts clear word divisions, and even goes beyond it in [ai l gou] and [bət θə piɡ

wud nt] with a blank in front of bound morphemes. Moreover, although Passy himself taught a variety of Southern English, it is interesting that he chose as editor to represent a rhotic accent. We argue, as shown by later discussions, that his transcription was not so much aimed at varieties of American English or conservative Northern English which were described as rhotic within the MP, but rather at providing a historical and panlectal (=cross-dialectal) notation in harmony with the standard orthography of English. But, then, the phonetic spelling offered was not really fully consistent. For instance, in other samples of the same period in the MP a long open back vowel is used in words like *past* or *master*, a feature of Received Pronunciation; which was definitely not a feature of American English or Northern rhotic accents as they were portrayed at the time or even today.

This policy was one that Sweet felt very uncomfortable with. What Sweet wanted was to make a clear distinction between (1) the nature of the international phonetic signs of the IPA, (2) much more readable spelling systems for each language not necessarily adhering to the one sound – one symbol principle, (3) pedagogical transcriptions which could vary according to purpose. The idea of a panlectal system was simply absent, each variety requiring its own system. Furthermore, he felt that a lot of symbols could be simplified. To give a short quote based on his variety of Southern English: “əz rɪgɑːdz ɪŋglɪʃ aɪ sɪj nou nɪjd fə ðɪ ʌnmɪjɪnɪŋ ənd ʌnrɪtəbl ʌ fər ə vɑːuəl ... whɪtʃ ɪz prækɪtɪklɪ ðə ʃɔːt əv ɑː” (i.e. ‘as regards English I see now no need for the unmeaning and unwritable ʌ for ə ... which is practically the short of ɑː’). Some of the changes suggested were ʌ → ɑ, ə → ɔ (possibly o), ø → ɵ, hw and ʌ → wh, as well as ø → e, and a → æ for French. Nor could he stop himself from making quite acid remarks on IPA symbols in so far as they departed from his own system of transcription.

Both Passy and Daniel Jones reacted very strongly against Sweet arguing that, despite all the respect and admiration they had for him, he was so to speak throwing out the baby with the bathwater [15, 16, 17]. Sweet’s death in 1912 may have liberated them since Passy and Jones ended up making even more drastic proposals concerning the transcription to be adopted in the MP. In the *Principles of the International Phonetic Association* of 1912, Passy and Jones as authors included towards the end of their presentation a brief example of ‘phonetic spelling’ for English [18]. The idea was that this form of writing should be easily understood and usable by speakers of all dialects. In reality, what Passy and Jones proposed was not truly phonetic spelling but rather a partially historical and panlectal system. It was based on the most conservative dialects, yielding much more abstract representations than the phonemic transcriptions they used for teaching. For instance, at the time when Passy and Jones were writing, the received Southern English pronunciation no longer maintained a distinction between words spelled with *wh-* and *w-*, such as *which* and *witch*. However, this contrast was still present in various English varieties spoken in the United States, Scotland, and Northern England, where it could be transcribed phonemically as /hwɪtʃ/ vs. /wɪtʃ/. Passy and Jones suggested maintaining this distinction in the so-called ‘phonetic’ spelling of English by consistently transcribing *which*, *when*, *why*, *what*, etc., with /hw/, while words like *witch*, *want*, and *wet* had a plain initial /w/. In the same vein, the transcription of English was that of a rhotic accent, since there is universal agreement that a coda ‘r’ in the English spelling (e.g. *bird*) corresponds to an historical sound. This approach mirrored conventional orthography, which often preserves letters corresponding to phonemes that have disappeared over time.

In the September–October 1913 issue of the MP [19], Jones further developed this idea in an article entitled ‘Phonetic Spelling of English’, and Passy supported him in an article a 1914 article entitled ‘Orthographe phonétique’ (*Phonetic orthography*), defending the same ideas as Jones just before the publication of the MP stopped because of the First World War.

Due to the interruption of the MP’s publication during the war, the issue remained unresolved. However, Passy renewed his efforts in the first post-war issue, the January–March

1923 edition of the MP. This time, rather than presenting a sample illustrating phonetic spelling, he published a three-page article entirely written in the new phonetic orthography, entitled *Vie Nouvelle* (*New Life*), transcribed as [vi: nuvelə] in the new phonetic spelling (as opposed to [vi nuvel] in all previous issues). To give the reader a better sense of the proposal, we reproduce the first paragraph below in the new system, followed by its transliteration, an English translation and our reconstruction of a phonetic transcription faithful to Passy's standard practice in the MP and *Les sons du français*:

New phonetic spelling: aprɛ ply də ʒit aŋ də mɔr aparɑ̃tə, lə mɛ:trə fɔnetikə rənɛ a la vi:! vwala la ʒwajøzə nuvelə, k avɛk no mɛʎœr swɛ, nuz aŋvwajɔŋ a tu le mɑ̃brɛ, aŋsjɛŋ e nuvo, də nɔtrə vjɛʎə e ʃɛrə asɔsjasjɔŋ. [20: 1]

Après plus de huit ans de mort apparente, Le Maître Phonétique renaît à la vie ! Voilà la joyeuse nouvelle, qu'avec nos meilleurs souhaits, nous envoyons à tous les membres anciens et nouveaux de notre vieille et chère association. ("After more than eight years of apparent death, Le Maître Phonétique is coming back to life! This is the joyful news that, with our best wishes, we send to all our members, old and new, of our old and dear Association.")

Passy's standard transcription style: aprɛ ply də ʒit ɑ̃ də mɔr aparɑ̃t, lə mɛ:tr fɔnetik rənɛ a la vi! vwala la ʒwajøz nuvel, k avɛk no mɛjœr swɛ, nuz ɑ̃vwajɔ̃ a tu le mɑ̃br, ɑ̃sjɛ̃ e nuvo, də nɔt vjɛj e ʃɛr asɔsjasjɔ̃.

The final part of the article (in translation here) clearly explained the overall project:

Jones and I have agreed that official articles should be written, not in proper transcription, but in *phonetic orthography* (underlined in the original), following the system outlined in the MP in September–October 1913, January–February, March–April, and July–August 1914. I am setting an example in this article... I already anticipate strong protests. Well then, here is a subject for interesting discussion: we are fully prepared to examine it in depth. Write to us with your thoughts, and we will respond—the issue is worth it." [20: 3]

Passy was not wrong to anticipate 'strong protests'. In the next issue of the MP (April–June 1923), Passy himself, in an article entitled *Orthographe ou transcription?* (*Spelling or Transcription?*), admitted that he had expected criticism but not "the tremendous outcry that arose," portraying him as "a traitor, a deserter, a renegade of the phonetic cause" [21: 10]. The issue included two strongly negative reactions in the 'Correspondence' section, one from a loyal member of the *Association*, Bessie Robson, who stated that the practical consequence of this innovation was that she "could no longer lend *Le Maître Phonétique* to her students' and the other from his long-time friend and collaborator, Danish scholar Elna Simonsen. She considered that Passy was "assassinating his own work" because, pedagogically, this orthography made no sense.

The study of subsequent issues of the journal reveals that Passy and Jones immediately abandoned the plan to publish all official articles in the new phonetic spelling, returning smoothly to conventional broad transcriptions very much in the Sweet tradition but maintaining now well established IPA symbols. What however remained of all these debates was the insistence on giving privilege to 'Romanic' symbols as Sweet called them along with a concerting effort to make many languages converge in the adoption of similar symbols from this Latin-based set.

6 Consequences and conclusion

The selective portions of history that we have examined in this paper show that the issue of spelling reform and its pedagogical implications has had an over-riding influence on the devising of the IPA. Despite sharp differences in approach, a consistent belief endured throughout most of the twentieth century—and even beyond—in the central importance of certain Latin-based symbols, which had been present from the very inception of the *Association Phonétique Internationale*. Even if Passy and Jones's more radical positions were defeated, a weakened position characterized by systematic use of 'broad' transcriptions (in Sweet's terminology), the retention of blanks, and standard punctuation, remained prevalent, and was implicitly regarded as an ideal form of spelling. If we wish to wrench away phonology and phonetics from techniques aimed at reducing languages to writing (a formulation used by Pike [22] as a definition of phonemics), we must become fully aware of the history of the symbols and the notational conventions we use.

We believe that the centrality of 'Romanic' symbols has had a considerable impact on the description of languages. A salient example of this has been the choice of /r/ for English, French (at the time) and a host of other languages (whether the variety uses e.g. [r], [r̥], [r̥̥], [r̥̥̥], [R], etc., as a main allophone), which can be shown to be highly problematic on general grounds [23]. But this is far from being the only example. Another case is the more or less systematic selection of /e/ and /o/ as default symbols whenever a language has only two mid vowels, a front unrounded one and a back rounded one. This selection may well be linked to the frequent assumption that if there are only two mid back and front vowels they are 'tense' (or mid-high) by default. But when we look at languages which have only vowels of this type (e.g. Spanish), the range of realizations between [e] and [ɛ] is very wide across phonotactic and dialectal contexts (see the contributions in [24]). A final illustration is the choice of /x/ as a symbol for a consonant which is in fact often more appropriately described as a uvular [χ] in many varieties of Arabic (an observation already made in [25]). These selections of Latin-based IPA symbols often give rise to panlectal descriptions giving the illusion of more symmetry than is actually attested, or constraints which require far more investigation, or the positing of historical processes for which the evidence is lacking. To avoid this, we need to make a fuller use of phonetic and psycholinguistic experimentation in addition to corpus-based phonological and phonetic analyses of languages. This is the only way to talk sensibly about phonological systems, their internal structures, their phonetic instantiations, and markedness statements about the languages of the world.

7 References

- [1] MACMAHON, M.K.C.: *The International Phonetic Association: The first 100 years*, *Journal of the International Phonetic Association*, 16, pp. 30–38, 1986.
- [2] SCHWEITZER, C. and C. DODANE: Du son au phonème : histoire des alphabets. In: C. DODANE and C. SCHWEITZER (eds): *Histoire de la description de la parole, de l'introspection à l'instrumentation*, pp. 189–18, Honoré Champion, Paris, 2021.
- [3] SCHWEITZER, C., C. DODANE and J. LAZAR: L'histoire des alphabets phonétiques du XVIIIe siècle jusqu'à l'API. In: *XXXIIe Journées d'Études sur la Parole, 4-8 juin 2018, Aix-en-Provence, France*, p 356-364. DOI: 10.21437/JEP.2018-41, 2018.
- [4] OHALA, J.J.: A brief historical survey of phonetic-phonological feature systems. In: L.-J. BOË and C.-E. VILAIN (eds): *Un siècle de phonétique expérimentale*, Lyon: ENS Éditions, Lyon, 2010. <https://doi.org/10.4000/books.enseditons.31100>.
- [5] LODWICK, F.: *An Essay towards an Universal Alphabet*, *Philosophical Transactions [Transactions of the Philosophical Society]*, no. 182, June 26, pp. 126–137, 1686.
- [6] DE BROSSES, C.: *Traité de la formation mécanique des langues, et des principes*

- physiques de l'étymologie*, 2 vol., Paris, Chez Saillant, Vincent, Desaint, Paris, 1765.
- [7] MONTMIGNON, J.-B.: *Système de prononciation figurée applicable à toutes les langues*, Royez, Paris, 1785.
- [8] DURAND, J. and C. LYCHE: *Paul Passy, un linguiste révolutionnaire. Réforme de l'orthographe, didactique des langues et alphabet phonétique international*, Lambert-Lucas, Limoges, 2024.
- [9] PASSY, P.: *La phonétique et ses applications*, *Supplement to Le Maître Phonétique*, 7(44), 1929.
- [10] PASSY, P.: *Le phonétisme au Congrès phonologique de Stockholm en 1886. Rapport présenté au Ministre de l'instruction publique*, Delagrave/Hachette, Paris, 1887.
- [11] BERGOUNIOUX, G.: La transcription des dialectes gallo-romans. L'exemple de l'alphabet de Rousselot. In L. SCHØSLER and J. HÄRMÄ (eds): *Actes du XXIXe Congrès international de linguistique et de philologie romanes*, vol 1, pp. 201–211, Éditions de Linguistique et de Philologie, Strasbourg, 2021.
- [12] SWEET, H.: *A Handbook of Phonetics, Including a Popular Exposition of the Principles of Spelling Reform*, Clarendon Press, Oxford 1877.
- [13] SWEET, H.: Normalised phonetic spelling, *Le Maître Phonétique* 21, pp. 71–73, 1906.
- [14] PASSY, P.: Anglais: the woman and her pig, *Le Maître Phonétique* 9(2), pp. 46–47, 1894.
- [15] PASSY, P.: Soyons conséquents, *Le Maître Phonétique* 21(12), pp. 127–128, 1906.
- [16] PASSY, P.: Écriture phonétique simplifiée, *Le Maître Phonétique* 22(3/4), pp. 35–37, 1907.
- [17] JONES, D.: Simplified phonetic spelling, *Le Maître Phonétique* 22(3/4), pp. 37–39, 1907.
- [18] PASSY, P. and D. JONES: *The Principles of the International Phonetic Association*. Supplement to *Le Maître Phonétique* 27, pp. 1–40, 1912.
- [19] JONES, D.: Phonetic spelling of English, *Le Maître Phonétique* 28, pp. 113–116, 1913.
- [20] PASSY, P.: Vie nouvelle, *Le Maître Phonétique* 1(38), pp 1–3, 1923.
- [21] PASSY, P.: Orthographe ou transcription?, *Le Maître Phonétique* 1(38), pp. 9–12, 1923.
- [22] PIKE, K.L.: *Phonemics. A Technique for Reducing Languages to Writing*, The University of Michigan Press, Ann Arbor, 1947.
- [23] ANSELME, R., F. PELLEGRINO and D. DEDIU: What's in the r? A review of the usage of the r symbol in the illustrations of the IPA, *Journal of the International Phonetic Association* 53(3), pp. 1–30, 2023.
- [24] GIL, J. and J. LLISTERRI (eds.): *Fonética y fonología descriptivas de la lengua española*, 2 Vol., Georgetown University Press, 2024.
- [25] LADEFOGED, P. and I. MADDIESON: *The Sounds of the World's Languages*, Blackwell, Oxford, 1996.

TRANSFER OF KNOWLEDGE ACROSS SPACES AND DISCIPLINES: THE CASE OF CZECH PHONETICS

Pavel Šturm

*Institute of Phonetics, Charles University, Prague, Czechia
pavel.sturm@ff.cuni.cz*

Abstract: This paper explores the transfer of knowledge across geographical and disciplinary boundaries in the development of Czech phonetics during the 20th century. It focuses on how international academic mobility – including study stays abroad and internships in Prague – contributed to the formation of a dense network of intellectual exchange. These personal connections were reinforced through correspondence, the exchange of books and offprints, and participation in international congresses, which functioned as critical hubs for the circulation of phonetic knowledge. Equally important was the interdisciplinary nature of early phonetic research, which drew from linguistics, acoustical physics, medicine or scientific cinematography. In turn, phonetics enriched adjacent domains such as speech therapy, language teaching, and dialectology. By examining the case of Czech phonetics, this study highlights the broader significance of such crossings in shaping the evolution of the phonetic sciences.

1 Introduction

Phonetics has long occupied a distinctive position at the intersection of multiple disciplines, including linguistics, medicine, physics, and psychology. As a science of spoken language, its development has often relied on interdisciplinary collaboration and the adaptation of tools and methods from other fields. Equally important has been the circulation of people, ideas, and practices across national and institutional boundaries, facilitated by academic mobility, international correspondence, and participation in dedicated conferences. As Goldsmith and Laks [9] convincingly demonstrate, the evolution of scientific fields is shaped not only by specific ideas, but more fundamentally by the contexts formed by individual scholars, institutional settings, and disciplinary traditions. This paper examines Czech phonetics as a case study of how such flows of knowledge – enabled by study stays abroad, internships in Prague, and sustained intellectual exchange in a strategic location between Western and Eastern Europe – contributed to the development of both the scientific foundations of phonetics and its broader applications in areas such as speech therapy.

2 Crossing spaces

2.1 Study stays abroad

Study or research stays at foreign institutions are a highly effective means of transferring knowledge. As the following examples illustrate, such exchanges often yield mutual benefits – for both the visiting scholar and the host institution. These experiences frequently occur early in a scholar's career, making the environment at the host institution particularly influential.

2.1.1 Josef Chlumský (1871–1939)

Chlumský began his academic training at Charles University in Prague, where he studied for two years before continuing his education abroad. His international studies included one semester in Strasbourg (then under German control), one in Berlin, and two semesters in Paris

during the academic year 1896/97. In Paris, he initially pursued Romance studies under Gaston Paris but soon developed a strong interest in phonetics, particularly through his engagement with abbé Rousselot (1846–1924), who had been teaching experimental phonetics at the Institute Catholique de Paris (ICP). During this period, Chlumský also witnessed the inauguration of the new phonetics laboratory at the Collège de France (CdF), which would soon gain international prominence. At the ICP, after completing two semesters of Rousselot's lectures and submitting an experimental phonetic thesis featuring kymographic recordings of Czech speech sounds (published as [3]), Chlumský was awarded the degree of *Maître en phonétique*.

However, a decisive moment in Chlumský's development was his second stay in Paris between 1910 and 1914. No longer a student, he collaborated closely with Rousselot, serving as his personal assistant and custodian of the laboratory at CdF. In this role, Chlumský tested new instruments, tutored Rousselot's students, conducted his own research, and contributed extensively to the *Revue de phonétique* (see [27] for an account of his numerous activities). Upon his return to Prague in 1914, he emerged as a dedicated and capable advocate of Rousselot's experimental methods, playing a key role in introducing and advancing them within the Czech phonetic tradition. Moreover, he retained a strong orientation toward French linguistic scholarship throughout his career.

2.1.2 Antonín Frinta (1884–1975)

Frinta spent three academic years studying in Prague and one year (1904–1905) in Paris, where he attended several prestigious institutions, including Université de Paris, Collège de France, and École pratique des hautes études. He studied under distinguished scholars such as Ferdinand Brunot, Antoine Meillet, Mario Roques, Jules Gilliéron, and most notably, Paul Passy (1859–1940). Under Passy's guidance, Frinta attended three key courses: a lecture on the diachronic development of French phonetics, a course on comparative phonetics of major European languages, and a seminar-style class dedicated to the analysis of phonetic texts.

Frinta adopted Passy's auditory phonetic methods and became an enthusiastic advocate of the International Phonetic Alphabet (IPA), which he applied extensively in his own work. He joined the Association in 1905 and actively contributed to its activities. Between 1906 and 1912, he published nine articles in *Le Maître Phonétique*, and in 1912 he was elected to the association's Council. Frinta developed a strong interest in pronunciation instruction (see 3.2.2). He promoted the direct method and fully embraced the association's principles in his teaching.

In 1919, Frinta was appointed as a reader at Charles University, where he began delivering phonetics lectures on a broad range of topics and languages, specializing in Slavic phonetics. His works and methods thus complemented those of Chlumský, who employed the experimental approach and focused on French and general phonetics.

2.1.3 Bohuslav Hála (1894–1970)

Hála, Chlumský's assistant in the phonetics laboratory, was sent to Strasbourg in the summer term of 1920 to improve his proficiency in French and to attend specialized courses in French phonetics alongside regular philology classes. His stay was arranged in cooperation with André Mazon, one of his lecturers during the period. This initiative indicates that Chlumský shared Frinta's belief in the importance of immersive, practical experience with a foreign language.

2.1.4 Přemysl Janota (1926–2008)

Janota belonged to a later generation of Czech phoneticians, educated after the Second World War. He was fortunate to study phonetics at Charles University before the communist regime came to power in 1948 and imposed strict limitations on Western influence. His early interest in Dutch was fostered by Bohumil Trnka, a member of the Prague Linguistic Circle. In 1947, Janota went to the University of Amsterdam, where he studied phonetics under the guidance of

Louise Kaiser (1891–1973), who had previously visited Chlumský’s laboratory during her own academic training and was personally acquainted with Hála. These existing connections made Janota’s placement in Amsterdam a natural choice. During his stay, Janota worked on his first experimental phonetic work, published in Dutch [13]. This marked the beginning of a lifelong friendship with Kaiser, as well as a sustained engagement with the Dutch language, its phonetic research traditions, and the country more broadly.

2.2 Internships in Prague

Between 1920 and 1939, as well as in the post-war period, the phonetics laboratory in Prague hosted a number of international interns and visiting scholars. They came from a range of countries – primarily France, but also Switzerland, Belgium, the Netherlands, the United Kingdom, the United States, China, Japan, Poland, Ukraine, or Yugoslavia (especially Serbia).

Most of them stayed only briefly, typically for a few days or weeks (cf. [25]). For instance, Ida Ward (1880–1949), a phonetics assistant at University College London, visited for a single day in June 1927. During her visit, her English pronunciation was recorded using a kymograph. Similarly, L. Kaiser, then an assistant at the Physiological Institute in Amsterdam, spent a week in Prague in 1923 to become familiar with the methods and research of the Prague phoneticians. During her stay, Chlumský made several kymographic recordings for her.

In contrast, some visitors stayed for several months or made repeated visits. One example is Ivan Pankevych (1887–1958), a dialectologist from Uzhhorod (present-day Ukraine), who visited the laboratory in 1922, 1923, and again in 1928. Initially, he came to learn about the available tools for dialect research, while later he used them for his own scholarly work. Among the many long-term interns, the following three individuals stand out.

2.2.1 Branko Miletić (1897–1983)

Miletić spent the academic year 1923/24 at the Prague laboratory. As one of the students of phonetics, he actively participated in all the laboratory’s classes and activities, gaining experience in physiological methods, kymographic recording, and auditory training. He attended the laboratory daily (excluding Sundays), working on his own research focused on the articulation of Serbian speech sounds – investigated through X-ray imaging – and Serbian melodic accents. This latter research formed the basis of his doctoral thesis. He remained at the laboratory for an additional year, during which he expanded his dissertation into a book [14] and completed his articulatory research for another publication [15]. Chlumský closely supported his work, reviewing and correcting Miletić’s manuscripts and research throughout his stay. Miletić returned to the laboratory several times during the 1920s and 1930s, usually for a month or two, to continue his research and make use of the laboratory’s equipment. During these visits, he also consulted on the technical requirements for establishing a phonetics laboratory in Belgrade.

2.2.2 Marguerite Durand (1904–1962)

Durand, an assistant at the Institut de Phonétique in Paris, was sent by Hubert Pernot to study under Chlumský during the summer semester of 1928/29. Her assigned project was to investigate French accentuation using experimental means. Under Chlumský’s guidance, she learned how to analyze accentuation in recordings, working with samples of her own speech, as well as those of Vendryès and two other native French speakers who happened to be in Prague at the time. She also actively participated in the broader activities of the class. The following year, Durand sent Chlumský a manuscript on French vowels, developed as a continuation of her seminar work from the previous summer. He provided detailed feedback, and the study was later published in *Revue de phonétique* [6].

2.2.3 Wu Tsung-Chi (1909–2010)

Wu, a visiting fellow from China (Chinese Academy of Science), arrived in Czechoslovakia in August 1957 and stayed until February 1958. His goal was to gain both practical and theoretical knowledge of modern experimental phonetic methods. During his stay, Wu visited 12 institutions across the country, with particular emphasis on the phonetics institute led by Hála. He met with 22 experts from various disciplines, studied general phonetics and acquired the basics of the Czech language. His training included a wide range of experimental techniques, such as spectral analysis, speech acoustics, synthesis, radiography, and other instrumental methods used in phonetic research. He also participated in fieldwork, assisting with dialect recordings. In addition to this phonetic training, he also immersed himself in the phonetic library.

2.3 Correspondence and exchange of publications

Beyond physical mobility, intellectual exchange through correspondence and the sharing of publications has long been central to academic collaboration. Letters served as a vital medium for discussing emerging ideas, critiquing methodologies and phonetic instruments, and staying informed about ongoing developments. Similarly, exchanging otherwise inaccessible journal articles and monographs between individual researchers supplemented phonetic libraries.

Chlumský corresponded with numerous prominent phoneticians and linguists, including Bogorodickij, Brunot, Chiba, Grammont, Kaiser, Mazon, Meillet, Menzerath, Millet, Panconcelli-Calzia, Pernot, Poirot, Rousselot, Scripture, and Vendryès, as well as with instrument makers such as Garten, Grandgent, Lioret, Montalbetti, and Struycken. In total, 72 international correspondents are documented (likely an underrepresentation), although the majority exchanged only a few letters. The most sustained correspondence, in terms of letter volume, was with Adrien Millet, Hubert Pernot, Georges Lote, Charles Ayer, Charles Bruneau, and E. W. Scripture. Millet and Pernot wrote over a hundred pages each, updating Chlumský on their work and on developments at several phonetic institutions in Paris (see [20] and [28]).

During the interwar period, Chlumský was not unique in this respect. Most active scholars across disciplines maintained substantial correspondence. For instance, Frinta's archive contains letters from hundreds of correspondents, including several phoneticians. However, most of the material reflects his broader interests in Slavic linguistics, literature, and culture. The collection remains largely unorganized, making comprehensive surveying difficult without browsing thousands of letters.

Frequently, books and offprints arrived along with the letters. E. W. Scripture, for instance, requested specific works from Chlumský and later asked to receive all his new publications. In return, Scripture enclosed his own writings, significantly enriching the holdings of the phonetic library in Prague, which now preserves many of his books and collected articles. This form of exchange was a common and valued means of acquiring scholarly literature.

2.4 Phonetic congresses

The International Congresses of Phonetic Sciences (ICPhS), officially established in 1932 following several smaller, locally organized events of more limited scope and disciplinary range, have since become a central part of professional life for many phoneticians, particularly those engaged in research. Czech phoneticians have been regular attendees and, incidentally, Prague holds the unique distinction of having hosted the congress twice, in 1967 and again in 2023.

The revival of the congress after World War II owes much to the persistent efforts of Hála, who advocated for its re-establishment for over a decade. It was partly thanks to him that the gap between the last pre-war congress in 1938 – which he personally attended – and the first post-war event in 1961 was ultimately bridged (see [26]). Hála understood well the value of such international gatherings, especially at a time when the communist regime severely restricted contact with researchers abroad. It is important to emphasize that the congresses were

more than venues for scientific exchange through lectures and discussions; above all, they provided opportunities to connect with colleagues and form new relationships, many of which were sustained through ongoing correspondence (see above).

3 Crossing disciplines

The previous sections focused on the transfer of (phonetic) knowledge across individuals, laboratories, and national boundaries, whether by means of physical relocation or remote correspondence. Section 3 will examine cases in which knowledge is transferred between disciplines or scientific environments. In this context, phonetics may be seen both as a field enriched by the insights and methodologies of other disciplines and as one that contributes its own knowledge and practices to them. Again, examples will be taken from Czech phonetics.

3.1 Enrichment of phonetics

3.1.1 Tuning forks and acoustic physics

Chlumský was captivated by Rudolph Koenig's Grand Tonometre, an elaborate and highly precise set of tuning forks acquired by Rousselot following Koenig's death. Chlumský set out to build a precision replica of the instrument which, while smaller in scope, featured 100 tuning forks covering a range from 16 Hz to 3,296 Hz (see [23] for further details).

Neither Rousselot nor Chlumský possessed the full range of skills required to construct a tonometer. The project demanded a specialized and interdisciplinary body of knowledge, combining artisanal craftsmanship, metallurgy and acoustic physics. Steel had to be selected with the same care as wood for musical instruments, and each fork had to be cast in a single piece, then painstakingly shaped and tuned for precision. Crucially, the project depended on a network of experts: R. Montalbetti played a key role in the production process, while a mechanic from Pirard et Coeurdevache was responsible for the finishing work. In 1921, Chlumský travelled to Paris to supervise this intricate process. He often collaborated with Rousselot and Millet, comparing the newly crafted forks with Koenig's originals to ensure accuracy. The calibration and testing of the forks took several months of daily work for just thirteen forks.

Tuning forks had practical applications in experimental phonetics. Chlumský employed them mostly as a pedagogical tool, training phonetic students and dialectologists to perceive subtle vowel distinctions. In the 1930s, Hála further utilized them for an extensive acoustic analysis of Czech vowels [10].

3.1.2 Medicine

The development of medicine in the 19th century was a key factor to the emergence of phonetics at the end of the century, when integrated with acoustic physics and, most importantly, linguistics [1]. Anatomical and physiological observations by 19th-century physicians and scholars significantly enriched phonetic descriptions of articulatory processes, grounding them in empirical evidence. Techniques such as vivisection, physiological experimentation, and measurement provided a new level of certainty to earlier speculative accounts. Instruments originally developed for physiological research were increasingly repurposed for the study of speech.

Phonetics was notably advanced by the contributions of two Czech physiologists: Jan Evangelista Purkyně (1787–1869) and Jan Nepomuk Czermak (1828–1873) (see [24]). Purkyně published a remarkably systematic physiological account of the articulation of speech sounds [19] two decades before Brücke's *Grundzüge der Physiologie* appeared [2]. Czermak, for his part, improved the laryngeal mirror, enabling direct observation of the vocal folds and effectively founding the field of clinical laryngoscopy. However, his work went beyond medical diagnostics: he also applied this instrument to the study of speech sounds produced at the glottis (e.g. [4]) and delivered popular lectures on the topic of “voice and speech” [5].

Collaboration between phonetics and medicine remained strong in the Czech lands. Emanuel Kovář (1861–1898), a linguist and self-taught phonetician, studied philology at the Faculty of Arts while also attending selected courses at the Faculty of Medicine. As a lecturer, he taught not only linguistics but also speech disorders to philology students. Deeply interested in the relationship between brain function and speech, Kovář was particularly drawn to emerging research on neurological impairments. His promising career was unfortunately cut short by suicide.

Later, Hála continued to cultivate strong ties to the medical community. He collaborated with Miloš Sovák (1905–1989), an otorhinolaryngologist specialized in phoniatrics who played a central role in the development of modern Czech speech therapy (see 3.2.1). Together, they authored a popular book on voice, speech, and hearing, which went through several editions [12]. Hála also maintained professional contacts with other notable figures in the field, including Raoul Husson (1901–1967).

3.1.3 Scientific cinematography

The desire to observe the activity of the vocal folds during phonation dates back centuries (cf. [16]). While Jan Nepomuk Czermak's improvements to the laryngeal mirror made direct visual examination of the vocal folds possible, the next major step involved capturing their rapid, cyclical movements on film. A significant advancement was achieved by Bohuslav Hála in 1928–1929 through his collaboration with Ludovít Honty, a specialist in scientific cinematography. Their work marked a breakthrough in visualizing glottal activity, combining high-speed cinematography with stroboscopic techniques, as detailed in [20]. The resulting footage, featuring Hála's own vocal folds, was presented to an international audience in 1930 during the Fourth International Congress of Logopaedics and Phoniatrics, held in Prague. The film was widely acclaimed, and a French-language report by Hála and Honty from 1931 remains a frequently cited source in the literature on phonation [11]. Panconcelli-Calzia recognized it as the first successful instance of capturing the vocal folds using high-speed cinematography [18].

3.1.4 Sound recording

One of the most essential technical advancements for phonetic research was the ability to record sound onto a physical medium and play it back for repeated analysis (auditory or instrumental). Early attempts, such as the use of the phonograph with wax cylinders, laid the groundwork, but it was the advent of more durable and reliable gramophone disc recordings that enabled broader distribution and systematic archiving. During the early 20th century, several speech archives were established across Europe, in cities such as Vienna, Paris, and Berlin. As early as 1907, Frinta participated in a phonographic training course organized by the University of Vienna and the Academy of Sciences, where he learned to record speech on specialized gramophone records for dialectological purposes, although he never really used that skill. It was not until the late 1920s that systematic efforts to archive Czech speech were undertaken.

Chlumský collaborated with French phoneticians and the Pathé recording company to initiate the first major wave of speech recordings in Czechoslovakia in 1929. The recorded material included dialect samples, public speeches, poetic recitations, or music. The endeavour was logistically demanding and required unprecedented sums of money. It was supported by the Czechoslovak Academy of Sciences and Arts and its Phonographic Commission, chaired by Chlumský. On the French side, the project was strongly supported by Hubert Pernot, a key collaborator. Together, they worked to overcome a range of technical challenges, including issues with the reliability of wax discs, which affected the consistency and quality of recordings.

3.2 Benefit of phonetics

This section turns to disciplines that have drawn on the findings of phonetics or adopted its methods to investigate spoken language. These fields often concentrate on specific contexts or populations – such as atypical speech, non-native speech, or dialectal variation – where the

application of phonetic analysis proves especially valuable. The final section addresses phonology, a discipline formally distinguished from phonetics during the interwar period, yet one that remains closely interconnected with it in both theory and practice.

3.2.1 Speech and language therapy (SLT)

Both Rousselot and Chlumský had significant ties to the treatment of atypical speech and hearing. Rousselot for instance diagnosed hearing loss by tonal testing with tuning forks, evaluating both the range and severity of impairment. In Prague, this connection was evident in the education and training of professionals in the field. Medical practitioners frequently visited the phonetic laboratory, where they learned to apply experimental phonetic research in clinical settings and address various speech disorders, including oesophageal speech.

A key figure in the development of Czech SLT was Hála, especially because of his contributions to the education and professional training of therapists. His influence extended partly from his phonetic lectures and published works. Hála's physiological studies on the voice had a lasting impact, informing later monographs on voice pathology [22]. His acoustic analyses of vowel characteristics, such as volume and carrying power, played a key role in auditory testing methods developed by his student, Přemysl Janota. Another prominent student, Miloš Sovák, became a key figure in SLT and advanced the field considerably.

Hála's primary influence, however, lay in his extensive extracurricular teaching. He organized numerous training sessions, courses, and workshops that combined practical applications with research-based insights. Hála delivered special lectures and demonstrations for teachers of the deaf and provided phonetics instruction for staff at the SLT Institute in Prague. He also participated in SLT training programs for both schoolteachers and preschool teachers, emphasizing the importance of early speech development and preventive phonetic care. One of his students, Karel Ohnesorg, went on to publish a foundational longitudinal study on children's speech development [17]. Hála's efforts were not confined to Prague, as he also travelled widely across the country, bringing phonetic education to rural areas and smaller institutions.

3.2.2 Teaching of foreign languages

The intersection of phonetics and language teaching was most clearly embodied in the work of Frinta. A dedicated secondary-school teacher of French, he retained this role even after his appointment as a lecturer at Charles University in 1919. He was also an active contributor to a local Czech journal *Lectures françaises*, where he published accounts of French pronunciation and examples of phonetic transcription under the column "The Phonetic Corner". Frinta was among the first educators to incorporate the gramophone into language instruction, using it to provide students with authentic pronunciation models (especially in terms of the prosody of speech). Chlumský likewise underscored the significance of this device for language teaching, even offering university lectures dedicated to the topic.

In 1909, Frinta published a monograph on the teaching of foreign pronunciation [7], aiming to improve prevailing practices at Czech schools. He criticized the limited attention given to pronunciation instruction and the widespread tolerance of strong Czech accents in foreign language classes. He held both students and teachers accountable for these shortcomings: students, for lacking motivation, and teachers, for failing to provide accurate pronunciation models and lacking a solid grounding in phonetics. Frinta emphasized the need for teachers to possess both practical and theoretical phonetic knowledge and recommended summer phonetics courses, which also allowed for language immersion in natural contexts. Most importantly, he advocated for the consistent use of phonetic transcription (IPA) as a central tool in pronunciation teaching. He argued that its benefits should be clearly explained to students, and that traditional orthography should be excluded from pronunciation instruction altogether.

3.2.3 Dialectology and language change

Frinta's work was instrumental in establishing the theoretical and methodological underpinnings of linguistic geography in the Czech context. He drew inspiration from scholars such as Jules Gilliéron and appreciated the meticulous, fieldwork-based approach of Rousselot [21]. Frinta perceived dialects as regionally varied forms of a national language, noting that while the boundaries between dialects were often fluid, it was possible to map precisely the distribution of individual pronunciation features (using *isophones*).

Phonetics played a central role in Frinta's dialectological research, as it enabled the direct observation and documentation of spoken language in real-world settings. He insisted that first-hand experience in local environments was essential, as questionnaires alone could not capture the nuance and variability of regional speech. His own fieldwork took him across various Slavic countries, including the Czech lands, Poland, and the Balkans, where he gathered empirical material that informed his studies of pronunciation and sound change. By focusing on phonetic phenomena such as the articulation of /v/, he was able to trace patterns of linguistic variation and link them to historical processes.

One of his most significant methodological contributions was the use of isophones to map the "colourful diversity" of spoken language. Layering numerous isophones onto a single map revealed complex patterns of variation that would have remained invisible in a purely theoretical approach. This method also allowed him to make well-supported reconstructions of older language stages. For instance, in his habilitation thesis [8], Frinta analyzed the current phonetic forms of labial consonants across Slavic regions to hypothesize the probable articulation of the Proto-Slavic /v/, which he concluded was bilabial. His insistence on beginning historical reconstruction from the detailed phonetic reality of contemporary speech exemplified his belief that dialectology must be grounded in the living language, not in written or standardised forms.

3.2.4 Phonology

In the interwar period, Prague was home to two prominent approaches to the study of spoken language. One was the *phonetic school* led by Chlumský, centred around his Laboratory of Experimental Phonetics, where Hála served as his assistant. Frinta, while deeply engaged in auditory phonetics, did not establish a formal school of his own but belonged to a loosely connected group of phoneticians and linguists working independently. In contrast, the second major body was the Prague Linguistic Circle (PLC), which developed a more comprehensive theoretical framework known as *functional structuralism*. Within this framework, the functional perspective was emphasized by Vilém Mathesius (1882–1945), the Circle's founder and leader, while the structural component was championed by N. S. Trubetzkoy (1890–1938) and Roman Jakobson (1896–1982), the principal architects of phonology.

From the outset, Trubetzkoy strongly insisted on a strict conceptual separation between phonetics and phonology, a stance articulated clearly in his early writings (e.g. [29]). This division, at times marked by rivalry or even tension, may have been necessary, as phonology was emerging as a distinct scientific discipline. Officially recognized as such at the First International Congress of Linguists in The Hague in 1928, phonology had to assert its legitimacy and differentiate itself from the already well-established field of experimental phonetics. The institutional imbalance at the time reflected this struggle: while there were two university chairs in phonetics (held by Chlumský and Frinta), there was no corresponding academic position for phonology.

Trubetzkoy's attitude toward phonetics was more nuanced than is sometimes assumed. While he famously argued for a strict separation between phonetics and phonology (going so far as to characterize phonetics as a biological science), his scepticism was primarily directed at the experimental phonetics practiced by figures like Panconcelli-Calzia. This type of phonetics, focused on instrumental and physiological measurement, struck Trubetzkoy as excessively empirical and disconnected from linguistic function. However, he did not reject phonetics as a

whole. On the contrary, he acknowledged the importance of auditory phonetics and phonetic transcription as essential steps in phonological analysis. In correspondence with Jakobson, Trubetzkoy even praised the British tradition for practical phonetic descriptions of exotic languages, noting that although it lacked theoretical sophistication, its findings could be easily translated into the language of phonology. His critique, then, was not of phonetics per se, but of a particular version of it that he saw as incompatible with structural analysis.

Mathesius, on the other hand, maintained a consistently positive relationship with phonetics throughout his career. A distinguished scholar of English, Mathesius was well-read in both experimental and auditory phonetics. He was particularly influenced by the practical orientation of scholars such as Sweet, Jespersen, Viëtor, and Passy, but he also intimately knew the experimental work of Rousselot, Meyer or Scripture. Mathesius was a member of the IPA, and he held the work of the Frinta in high esteem. Though primarily a linguist and a phonologist, Mathesius clearly saw phonetic analysis as an indispensable tool for understanding language structure.

Chlumský held a markedly different view on the relationship between phonetics and phonology. He regarded phonology not as an independent discipline but as a subfield of phonetics – specifically, a third dimension complementing the articulatory and acoustic aspects of speech sounds. From this perspective, he argued that phonology offered nothing fundamentally new, and that phonetics already encompassed a more complete and comprehensive study of speech. However, Chlumský misunderstood the aims and scope of phonology, interpreting it primarily as the psychology of speech sounds rather than as the analysis of sounds within the structure of a language system. As a result, central phonological concepts such as the phoneme, functional opposition, and the notion of phoneme realization were largely foreign to his thinking.

4 Conclusions

Czech phonetics was used as a case study to illustrate how knowledge circulated within the early phonetic community in Europe. A key factor in the development of individual phoneticians was their early exposure to phonetics in international settings, where specific institutions, people, and contexts shaped their future careers. At the same time, the Prague laboratory served as an important hub for training foreign students, particularly from Central and Eastern Europe, offering hands-on experience in experimental phonetics. Personal correspondence between scholars often sustained further long-term intellectual exchange and may have been more influential than the relatively infrequent phonetic conferences.

Another central theme was the inherently interdisciplinary nature of phonetics. The field has long drawn on tools and concepts developed in other domains – such as tuning forks from acoustical physics or kymographs from medical research – and repurposed them for the study of speech. However, phonetics was also a source of inspiration for other disciplines that worked with human speech. This was most evident in dialectological studies and phonology, which needed precise phonetic descriptions, or in speech therapy and foreign language instruction, where phonetic knowledge and tools were crucial for successful acquisition.

5 Acknowledgments

This publication was supported by the Cooperatio program of Charles University, Linguistics research area, implemented at the Faculty of Arts of Charles University.

6 References

- [1] ASHBY M. & J. MCELVENNY. 2022. The emergence of phonetics as a field. In *Interviews in the History of Linguistics*, 51–60. Language Science Press.
- [2] BRÜCKE, E. 1856. *Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Taubstummlehrer*. Carl Gerold's Sohn.

- [3] CHLUMSKÝ, J. 1902. Analyse du courant d'air phonateur en tchèque. *La Parole* 4. 130–143, 395–406, 483–496, 665–683, 736–741.
- [4] CZERMAK, J. N. 1879. Über den Spiritus asper und lenis, und über die Flüsterstimme, nebst Bemerkungen zur phonetischen Transscription der Kehlkopflaute. In *Gesammelte Schriften I*, 750–766. Wilhelm Engelmann.
- [5] CZERMAK, J. N. 1879. Stimme und Sprache. In *Gesammelte Schriften II*, 60–104. Wilhelm Engelmann.
- [6] DURAND, M. 1929. Étude sur les phonèmes postérieurs dans une articulation parisienne. *Revue de phonétique* 6. 241–256.
- [7] FRINTA, A. 1909. O učení cizí výslovnosti [On teaching foreign pronunciation]. In *Pátá výroční zpráva reálky v Mladé Boleslavi*. Hejda a Zbroj.
- [8] FRINTA, A. 1916. *Fonetická povaha a historický vývoj souhlásky „v“ ve slovanštině [Phonetic Nature and Historical Development of “v” in Slavic]*. Česká akademie.
- [9] GOLDSMITH, J. & B. LAKS. 2019. *Battle in the Mind Fields*. University of Chicago Press.
- [10] HÁLA, B. 1941. *Akustická podstata samohlásek [Acoustical Basis of Vowels]*. ČAVU.
- [11] HÁLA, B. & L. HONTY. 1931. La cinématographie des cordes vocales à l'aide du stroboscope et de la grande vitesse. *Otolaryngologia Slavica* 3. 1–12.
- [12] HÁLA, B. & M. SOVÁK. 1941. *Hlas, řeč, sluch [Voice, Speech, Hearing]*. Česká grafická unie.
- [13] JANOTA, P. 1947. *Onderzoek betreffende de kwantiteit van de Tsjechische en Nederlandse klinkers [Research on the Quantity of Czech and Dutch Vowels]*. University of Amsterdam.
- [14] MILETIĆ, B. 1926. *O srbo-chorvatských intonacích v nářečí štokavském [On Serbo-Croatian Intonations in the Štokavian Dialect]*. FF UK.
- [15] MILETIĆ, B. 1933. *Izgovor srpskohrvatskih glasova [Pronunciation of Serbo-Croatian Sounds]*. Srpska kraljevska akademija.
- [16] MOORE, P. 1991. A short history of laryngeal investigation. *Journal of Voice* 5(3). 266–281.
- [17] OHNESORG, K. 1948. *Fonetická studie o dětské řeči [A Phonetic Study on Child's Speech]*. FF UK.
- [18] PANCONCELLI-CALZIA, G. 1931. [Part of Bibliographia Phonetica]. *VOX* 17(1). 77–78.
- [19] PURKYNĚ, J. E. 1835 [1958]. *Pojednání přírodoslovné o mluvě lidské [A Natural Science Treatise on the Human Speech]*. In *Jan Evangelista Purkyně – Opera Omnia (Volume XII)*, 89–113. ČSAV.
- [20] RIALLAND, A., P. ŠTURM & C. FOUGERON. 2024. Rousselot's succession through Pernot and Millet's correspondence to Chlumský. In *Proc. Sixth International Workshop on the History of Speech Communication Research*, Budapest, 25–34. TUDpress.
- [21] ROUSSELOT, P.-J. 1891. *Les modifications phonétiques du langage étudiées dans le patois d'une famille de Cellesrouin (Charente)*. H. Welter.
- [22] SOVÁK, M. 1945. *Stroboskopický výzkum hlasové pathologie [Stroboscopic Research in Voice Pathology]*. ČAVU.
- [23] ŠTURM, P. 2015. The Prague historical collection of tuning forks: A surviving replica of the Koenig tonometre. In *Proc. First International Workshop on the History of Speech Communication Research*, Dresden, 95–105. TUDpress.
- [24] ŠTURM, P. 2019. The contribution of Czech phonetics to laryngeal investigation. In *Proc. 19th International Congress of Phonetic Sciences*, Melbourne, 1903–1907. ASSTA.
- [25] ŠTURM, P. 2022. The Laboratory of Experimental Phonetics in Prague between 1919 and 1939: Research practices and student instruction. In *Proc. Fifth International Workshop on the History of Speech Communication Research*, Porto, 75–85. TUDpress.
- [26] ŠTURM, P. 2023. The 6th ICPHS in Prague and the organized life of phonetics. In *Proc. 20th International Congress of Phonetic Sciences*, Prague, 3231–3235. Guarant International.
- [27] ŠTURM, P. 2025. From Paris to Prague: Joseph Chlumský's enduring commitment to Rousselot's experimental phonetics. In *Actes du Colloque en hommage à l'Abbé Rousselot*. Société de linguistique de Paris.
- [28] ŠTURM, P., A. RIALLAND & C. FOUGERON. 2024. Pernot, Scripture and Millet: Different ideas for running a journal of experimental phonetics in the early 20th century. In *Proc. Sixth International Workshop on the History of Speech Communication Research*, Budapest, 35–45. TUDpress.
- [29] TRUBETZKOY, N. S. 1929. Zur allgemeinen Theorie der phonologischen Vokalsysteme. *Travaux du Cercle Linguistique de Prague* 1. 39–67.

EARLY TRANSCRIPTIONS OF ENGLISH AND FRENCH IN THE INTERNATIONAL PHONETIC ASSOCIATION: THE WEIGHT OF SPELLING REFORM ON THE BIRTH OF THE IPA

Jacques Durand¹, Chantal Lyche²

¹CLLE, Université Toulouse-Jean Jaurès, France, ²University of Oslo, Norway
jacques.durand@univ-tlse2.fr

Abstract: The establishment of the International Phonetic Association and the creation of its alphabet (the IPA) attest to the fact that scientific advancement emerges from collective collaboration and sustained debate. The development of the IPA is closely intertwined with progress in the fields of phonology and phonetics. However, the founders of the Association—such as Paul Passy, Otto Jespersen, and Wilhelm Viëtor—alongside influential scholars like Isaac Pitman, Alexander Ellis, and Henry Sweet, were also ardent advocates of spelling reform. We examine the extent to which this reformist agenda influenced the evolution of the IPA, with particular focus on debates within the *Maître Phonétique* concerning the concepts of ‘phonetic transcription’ and ‘phonetic spelling.’ We show that spelling reform was consistently viewed as a core concern and constituted a persistent, underlying thread in the evolution of the IPA. We argue that much of the debate between Sweet, Passy and Jones in *Le Maître Phonétique* had no winner as such. However, the adoption of Sweet’s ‘broad’ transcriptions preferentially based on Latin-based symbols identically used across languages led to a confusion between the task of phonology/phonetics, the devising of spelling systems and the pedagogical applications of the IPA.

1 Introduction

The establishment of the International Phonetic Association and the creation of its alphabet (the IPA) attest to the fact that scientific advancement emerges from collective collaboration and sustained debate. The development of the IPA is closely intertwined with progress in the fields of phonology and phonetics. However, the founders of the Association—such as Paul Passy, Otto Jespersen, and Wilhelm Viëtor—alongside influential scholars like Isaac Pitman, Alexander Ellis, and Henry Sweet, were also ardent advocates of spelling reform. After a brief examination of this historical background, we study the foundation of the Association and show that there was much ambiguity between the devising of an international phonetic alphabet for scientific purposes, spelling reform and pedagogical applications. The controversies among various figures—particularly between Sweet and Passy/Jones within *Le Maître Phonétique*—had a profound impact on the development of the IPA. The push for a quasi-universal use of Latin-based symbols to represent similar, yet often distinct, sounds across languages still echoes in contemporary practice. Moreover, the adoption of ‘broad’ transcriptions featuring word blanks and standard punctuation has long contributed to a blurring of the line between phonological representations and orthographic systems.

2 Historical background

The development of the International Phonetic Alphabet (henceforth IPA) was a collaborative endeavour, it was “hundreds of people around the globe reading and discussing every month topics in phonetics and language-teaching.” [1] Nonetheless, it was Paul Passy (1859–1940)

who played a pivotal role in steering the project forward. His unwavering commitment culminated in the establishment of an organization that would later become the *Association Phonétique Internationale* in 1897 and, subsequently, the International Phonetic Association in 1970.

The concept of a universally applicable phonetic alphabet was by no means novel in the late 19th century. Historical evidence indicates that the idea of one-to-one correspondences between phonemes and graphemes has ancient origins, frequently underlying early alphabetic writing systems [2], [3]. Universal alphabets characterizing speech sounds in terms of phonetic features were already being devised in the 17th and 18th centuries [4]. Notably, in 1686, Francis Lodwick (1619–1694), a Flemish scholar, published *An Essay toward a Universal Alphabet* [5]. In this work, he articulated a visionary rationale for such an alphabet. Lodwick proposed that a universal phonetic system could facilitate children's acquisition of foreign pronunciations and enable the accurate representation of any language's pronunciation. He wrote:

It will also be useful to perpetuate the true Sounds of any Language, and serve as a Standard thereof to after Ages: for if all the single Sounds expressable, be here Collected and Characterized, *And that no one Character have more than one Sound, nor any one Sound be expressed by more than one Character*; it cannot fall out that any Character should be falsly pronounced, but it will soon be discovered; for this false Sound he giveth it, must be the true Sound of some other Letter of this Alphabet: and so none can Erre herein, but he that wilfully or carelesly will do it. [5, p. 3, original spelling]

In 18th-century France, two additional projects merit mention. Charles de Brosses (1709–1777) devised an “organic universal alphabet” primarily intended for cross-linguistic comparisons [6], while Jean-Baptiste Montmignon (1737–1824) developed a “figurative alphabet applicable to all languages” [7], echoing many of Lodwick’s objectives [2], [3].

Nineteenth-century initiatives to create a universal, phonetically based alphabet thus drew upon this long-standing intellectual tradition concerned with the systematic transcription of human language. Among the newly developed systems, those that remained closest to familiar European alphabets tended to be the most widely accepted. Two prominent examples illustrate this trend. François Volney advanced several proposals and, in 1804, published the influential *L'alphabet européen appliqué aux langues latines*. He also left a bequest of 24,000 francs to the *Institut de France*, instructing that it be awarded to any individual capable of devising a “harmonic alphabet” based on Latin script—one that would bring coherence to the fragmented orthographic practices of the time. Another significant contribution came from the German Egyptologist Carl Richard Lepsius, who published his *Standard Alphabet* in 1854 (German edition), followed by an English edition in 1855. Intended as a uniform orthographic system based on ‘European’ characters, the alphabet was designed for use with both unwritten languages and transcriptions of non-European scripts. Despite its scientific rigor and success—particularly in Africa—the system’s reliance on an extensive set of diacritics discouraged widespread adoption among users and publishers.

Thus, while the idea of a universal phonetic alphabet was very much in keeping with the intellectual spirit of the 19th century, its broader success depended not only on conceptual soundness, but also on the presence of influential advocates capable of demonstrating its feasibility to a larger scholarly and pedagogical community. Paul Passy would take up this challenge with unwavering dedication, joined by international colleagues who shared his conviction in the necessity of such an alphabet.

3 The birth of the IPA

Let us briefly recall that Passy, as early as 1882-1883, had published *L'Anglais parlé, ou Méthode pratique pour apprendre à parler, à comprendre et à lire en anglais sans maître* (154 pages in two volumes) under the imprint of Eizak Pitman, Fonetic Institute [sic]. The originality of this publication lies not so much in the contents of the lessons but in the complete adoption of a phonetic transcription based on Pitman's 'phonotypic alphabet' for all the English examples instead of the standard spelling. This choice marks the first entry by Passy into the world of modern phonetics. Passy, determined to improve foreign language teaching, strongly believed that learners would quickly acquire an adequate pronunciation if presented with texts in phonetic spelling rather than in standard orthography. Passy created on 2 January 1886 the *Association phonétique des professeurs d'anglais*, soon after it was enlarged to include all language teachers and became *Dhi Fonètik Tîtcierz' Asóciécon* with its journal *Dhi Fonètik Tîtcer* first published in May 1886. The International Phonetic Alphabet (IPA) then began modestly—compared to Lodwick's universal alphabet—, with a relatively limited initial objective: the transcription of English. Its broader development was catalysed by Paul Passy's participation in the Stockholm philological meeting in August 1886, during which the *Quousque Tandem Scandinavian Association for the Reform of Language Teaching* was established. The *Association's* first proposal emphasized the primacy of oral language in language instruction, advocating for the early introduction of phonetic transcriptions of texts rather than reliance on conventional orthography. Passy, along with other proponents like Otto Jespersen, sought to extend this approach to all languages. However, in order to build broader consensus, the initial implementation was limited to languages characterized by a significant divergence between standard spelling and actual pronunciation. The connection between the *Association* and language teaching remained robust for a long period. In January 1889, *The Phonetic Teacher* was renamed *Le Maître Phonétique*, and at the same time, *The Phonetic Teacher's Association* became *L'Association Phonétique des Professeurs de Langues Vivantes*, with Wilhelm Viëtor serving as its first president [8]. In 1897, the *Association* expanded its scope and adopted the name *Association Phonétique Internationale*, reflecting its increasingly international membership, which by then included 865 members from a diverse range of countries. The most recent structural revision occurred in 1970, when the organization formally became the *International Phonetic Association*, and *Le Maître Phonétique* was succeeded by the *Journal of the International Phonetic Association*.

Paul Passy consistently placed language instruction at the center of his vision for phonetics. In a 1929 analysis of the main applications of phonetic science, he emphasized the importance of foreign language instruction concluding: "One could nearly say that without Phonetics, modern languages seem dead; with Phonetics, dead languages seem alive" [9: 9]. His international collaborators (such as Jespersen, Viëtor, and Lundell, to name a few) were all prominent phoneticians and language educators who largely viewed the phonetic alphabet as a pedagogical tool. Jespersen, for instance, had suggested to Passy the renaming of the organization to *L'Association Phonétique des Professeurs de Langues Vivantes* [10: 12]: "Would it not be preferable to have an Association of teachers of *all modern* languages? The same considerations that plead in favour of the phonetic method for the teaching of English, indeed make it a necessity, apply equally well to the teaching of French, Russian, Danish, etc., etc."

However, language pedagogy was not the sole impetus behind Passy's foundational work that eventually led to the development of the International Phonetic Alphabet. Passy's ambition extended beyond this practical application and also envisaged the phonetic alphabet as a possible alternative to traditional orthography although the first issue of *The Phonetic*

Teacher (May 1886) asserts the opposite¹: “There are lots of spelling reform papers already, what’s the use of having one more? True, reader; but this isn’t going to be “one more”. We have nothing to do with spelling reform; individually, we may favour it, but as a body, we know nothing about it. Listen and I’ll tell you who and what we are, and why we start a paper.” Despite this disclaimer, the fact that both *The Phonetic Teacher* and *Le Maître Phonétique* (henceforth MP) were published entirely in phonetic transcription underscores this broader objective. It is therefore noteworthy that, although the IPA was regarded as useful for linguistic description, this was always a secondary consideration for Passy. In contrast, the eminent phonetician Pierre-Jean Rousset (1846–1924)—who notably refused to adopt the IPA [8]—was motivated solely by linguistic concerns when designing his own alphabet. However, with its 288 possible vowels, his system proved too complex and was soon forgotten [11].

4 The spelling issue

Passy's experience in language teaching led him to be the driving force behind two closely related endeavours: a spelling reform and the development of a phonetic alphabet. In doing so, he joined a broader movement of the time that advocated for the reform of standard spelling.

His work was significantly influenced by British pioneers, particularly Isaac Pitman (1813–1897) and Henry Sweet (1845–1912). Pitman, himself an educator, recognized the detrimental effects of conventional English orthography and began reimagining English transcription based on its phonological structure. This initiative resulted in the creation of a phonetic shorthand system, published in 1837 under the title *Stenographic Sound-Hand*. Subsequently, Pitman collaborated with Alexander Ellis (1814–1890) to develop a “phonotypic alphabet”, which proved to be a major influence on Passy. In particular, this alphabet introduced several symbols outside the traditional Latin set—some of which would later be incorporated into the IPA with their modern phonetic values. These include [ʃ] and [ʒ] for the sounds in *vicious* and *vision*, respectively, as well as [ŋ] for the <ng> in *sing*. Other symbols—such as precursors to the modern IPA [ɛ], [ɣ], and [ʊ]—were also present, although they initially had different phonetic values. A variant of this alphabet was devised by Passy and appeared in the first issue of *The Phonetic Teacher* (May 1886), which opened with the line: “Dher are lots of spelin reform pépers ôlredi; hwot’s di yûs ov havin wœn mór?”

Sweet followed in Pitman’s footsteps in his pursuit of a standard orthography more closely aligned with actual pronunciation. His first major work, *A Handbook of Phonetics*, published in 1877, carried the subtitle *A Popular Exposition of the Principles of Spelling Reform*, reflecting this commitment. The *Handbook* had a profound influence on Passy and played a foundational role in shaping the principles underlying the IPA. Sweet, who joined the *Association Phonétique* in 1886 and was appointed as Honorary President of the *Phonetic Teacher’s Association* in January 1888, fully supported the younger generation’s initiative to introduce phonetic texts in the classroom before standard orthography. In his *Handbook*, he also advocates for the need for spelling reforms: “Of course the orthographies of most of the Continental languages require reform as well as English; French, especially, most urgently demands a thorough change. Indeed, there is no reason why foreigners should not learn French on a phonetic system, leaving the present French spelling to be acquired afterwards, even if the French themselves do not inaugurate a reform.” [12: 198] In his exploration of several transcription systems (organic, Romic, glossic) he favours the Romic (or Romanic), Latin-based, which the IPA is partially based on.

¹ The quote is given here in standard orthography.

Sweet defended the necessity of separating two types of transcription, a ‘broad’ and a ‘narrow’ one, which anticipated the modern concept of phoneme: “Again, among the vowels of any one language we must distinguish between those differences which are *distinctive* [underlined in the original], that is to which differences of meaning correspond, and those which are not. Thus, the first element of the diphthongs in ‘by’ and ‘out’ vary considerably: some people sound them broad as in ‘father,’ some thin, as in ‘man,’ with various intermediate sounds. And yet the meaning of the words remains unchanged. The distinction between the vowels of ‘men’ and ‘man,’ on the other hand, though really slighter than that of the different pronunciations of ‘by’ and ‘out,’ is a distinctive one” [12: 182]. This position is strongly supported by Passy and systematically applied in the MP and in his analyses of French [8]. However, Sweet in his objective to “make phonetic spelling less repellent to the general public” [13: 73] made proposals that Passy and Jones strongly refuted.

5 Debates on notation

The elements sketched above provide the broad background against which debates on notation took place within the *Maître Phonétique*. Passy as general editor was instrumental in determining which direction the debate would take. From more or less the beginning, the line he followed was to advocate the same phonetic notation both as the medium of communication between members and as a pedagogical tool providing reading material in the classroom. The conflation of these two functions explains ambiguities that were never solved. For instance, whereas the word was considered not to have any phonetic reality, transcribed utterances preserved blank spaces between words going even beyond traditional divisions as we shall see.

To demonstrate our point, let us take an example from the ‘Partie des élèves’ (Learners’ Corner) from 1894 [14] which uses the transcription adopted by Passy in all his contributions in English to this journal, preceded by our transliteration. This passage provides an illustration of the kind of material often included by Passy for beginners. We will not focus on the advantages and possible drawbacks of such texts but note that, as time went on, fewer and fewer pedagogical samples were submitted to the MP, a problem regularly bemoaned by Passy.

The woman and her pig

An old woman was sweeping her house one morning, and she found a little crooked sixpence. “What shall I do”, said she, with this little sixpence? I’ll go to market, and buy a little pig, and come home tonight to cook my apple-dumplings.”

So she went to market, and bought a little piggy-wiggy. As she was coming home, she came to a stile; the pig would’nt go over the stile. She said “pig, pig, go over the stile, or I shan’t get home tonight, to cook my apple-dumplings, oh.” – but the pig wud’nt.

ðə wʊmən ənd ər piɡ

ən ould wʊmən wəz swi:pɪŋ ər haus wʌn mɔ:rnɪŋ, ən ʃi faʊnd ə litl krukɪd sɪkspəns. “hwɒt ʃl ai duw”, sed ʃi, “wið ðis litl sɪkspəns? ai l gou tə mɑ:rkɪt, ən bai ə litl piɡ, ən kʌm hoʊm tənaɪt tə kuk mai æpl-dʌmplɪŋ.”

sou ʃi went tə mɑ:rkɪt, ən bɔ:t ə litl piɡi-wiɡi. æz ʃi wəz kʌmiŋ hoʊm, ʃi keɪm tə ə stɑɪl; ðə piɡ wud nt gou ɔʊvəɪ ðə stɑɪl. ʃi sed “piɡ, piɡ, gou ɔʊvəɪ ðə stɑɪl, ər ai ʃənt get hoʊm tənaɪt, tə kuk mai æpl-dʌmplɪŋz, ɔʊ.” – bət θə piɡ wud nt.

From a phonetic point of view, a first observation is that by 1894 the symbols were more or less identical with the ones used today. We can immediately see that this transcription of English adopts clear word divisions, and even goes beyond it in [ai l gou] and [bət θə piɡ

wud nt] with a blank in front of bound morphemes. Moreover, although Passy himself taught a variety of Southern English, it is interesting that he chose as editor to represent a rhotic accent. We argue, as shown by later discussions, that his transcription was not so much aimed at varieties of American English or conservative Northern English which were described as rhotic within the MP, but rather at providing a historical and panlectal (=cross-dialectal) notation in harmony with the standard orthography of English. But, then, the phonetic spelling offered was not really fully consistent. For instance, in other samples of the same period in the MP a long open back vowel is used in words like *past* or *master*, a feature of Received Pronunciation; which was definitely not a feature of American English or Northern rhotic accents as they were portrayed at the time or even today.

This policy was one that Sweet felt very uncomfortable with. What Sweet wanted was to make a clear distinction between (1) the nature of the international phonetic signs of the IPA, (2) much more readable spelling systems for each language not necessarily adhering to the one sound – one symbol principle, (3) pedagogical transcriptions which could vary according to purpose. The idea of a panlectal system was simply absent, each variety requiring its own system. Furthermore, he felt that a lot of symbols could be simplified. To give a short quote based on his variety of Southern English: “əz rɪgɑːdz ɪŋglɪʃ aɪ sɪj nou nɪjd fə ðɪ ʌnmɪjɪnɪŋ ənd ʌnrɪtəbl ʌ fər ə vɑːuəl ... whɪtʃ ɪz prækɪtɪklɪ ðə ʃɔːt əv ɑː” (i.e. ‘as regards English I see now no need for the unmeaning and unwritable ʌ for ə ... which is practically the short of ɑː’). Some of the changes suggested were ʌ → ɑ, ə → ɔ (possibly o), ø → ɵ, hw and ʌ → wh, as well as ø → e, and a → æ for French. Nor could he stop himself from making quite acid remarks on IPA symbols in so far as they departed from his own system of transcription.

Both Passy and Daniel Jones reacted very strongly against Sweet arguing that, despite all the respect and admiration they had for him, he was so to speak throwing out the baby with the bathwater [15, 16, 17]. Sweet’s death in 1912 may have liberated them since Passy and Jones ended up making even more drastic proposals concerning the transcription to be adopted in the MP. In the *Principles of the International Phonetic Association* of 1912, Passy and Jones as authors included towards the end of their presentation a brief example of ‘phonetic spelling’ for English [18]. The idea was that this form of writing should be easily understood and usable by speakers of all dialects. In reality, what Passy and Jones proposed was not truly phonetic spelling but rather a partially historical and panlectal system. It was based on the most conservative dialects, yielding much more abstract representations than the phonemic transcriptions they used for teaching. For instance, at the time when Passy and Jones were writing, the received Southern English pronunciation no longer maintained a distinction between words spelled with *wh-* and *w-*, such as *which* and *witch*. However, this contrast was still present in various English varieties spoken in the United States, Scotland, and Northern England, where it could be transcribed phonemically as /hwɪtʃ/ vs. /wɪtʃ/. Passy and Jones suggested maintaining this distinction in the so-called ‘phonetic’ spelling of English by consistently transcribing *which*, *when*, *why*, *what*, etc., with /hw/, while words like *witch*, *want*, and *wet* had a plain initial /w/. In the same vein, the transcription of English was that of a rhotic accent, since there is universal agreement that a coda ‘r’ in the English spelling (e.g. *bird*) corresponds to an historical sound. This approach mirrored conventional orthography, which often preserves letters corresponding to phonemes that have disappeared over time.

In the September–October 1913 issue of the MP [19], Jones further developed this idea in an article entitled ‘Phonetic Spelling of English’, and Passy supported him in an article a 1914 article entitled ‘Orthographe phonétique’ (*Phonetic orthography*), defending the same ideas as Jones just before the publication of the MP stopped because of the First World War.

Due to the interruption of the MP’s publication during the war, the issue remained unresolved. However, Passy renewed his efforts in the first post-war issue, the January–March

1923 edition of the MP. This time, rather than presenting a sample illustrating phonetic spelling, he published a three-page article entirely written in the new phonetic orthography, entitled *Vie Nouvelle* (*New Life*), transcribed as [vi: nuvelə] in the new phonetic spelling (as opposed to [vi nuvel] in all previous issues). To give the reader a better sense of the proposal, we reproduce the first paragraph below in the new system, followed by its transliteration, an English translation and our reconstruction of a phonetic transcription faithful to Passy's standard practice in the MP and *Les sons du français*:

New phonetic spelling: aprɛ ply də ʒit aŋ də mɔr aparɑ̃tə, lə mɛ:trə fɔnetikə rənɛ a la vi:! vwala la ʒwajøzə nuvelə, k avɛk no mɛʎœr swɛ, nuz aŋvwajɔŋ a tu le mɑ̃brɛ, aŋsjɛŋ e nuvo, də nɔtrə vjɛʎə e ʃɛrə asɔsjasjɔŋ. [20: 1]

Après plus de huit ans de mort apparente, Le Maître Phonétique renaît à la vie ! Voilà la joyeuse nouvelle, qu'avec nos meilleurs souhaits, nous envoyons à tous les membres anciens et nouveaux de notre vieille et chère association. ("After more than eight years of apparent death, Le Maître Phonétique is coming back to life! This is the joyful news that, with our best wishes, we send to all our members, old and new, of our old and dear Association.")

Passy's standard transcription style: aprɛ ply də ʒit ɑ̃ də mɔr aparɑ̃t, lə mɛ:tr fɔnetik rənɛ a la vi! vwala la ʒwajøz nuvel, k avɛk no mɛjœr swɛ, nuz ɑ̃vwajɔ̃ a tu le mɑ̃br, ɑ̃sjɛ̃ e nuvo, də nɔt vjɛj e ʃɛr asɔsjasjɔ̃.

The final part of the article (in translation here) clearly explained the overall project:

Jones and I have agreed that official articles should be written, not in proper transcription, but in *phonetic orthography* (underlined in the original), following the system outlined in the MP in September–October 1913, January–February, March–April, and July–August 1914. I am setting an example in this article... I already anticipate strong protests. Well then, here is a subject for interesting discussion: we are fully prepared to examine it in depth. Write to us with your thoughts, and we will respond—the issue is worth it." [20: 3]

Passy was not wrong to anticipate 'strong protests'. In the next issue of the MP (April–June 1923), Passy himself, in an article entitled *Orthographe ou transcription? (Spelling or Transcription?)*, admitted that he had expected criticism but not "the tremendous outcry that arose," portraying him as "a traitor, a deserter, a renegade of the phonetic cause" [21: 10]. The issue included two strongly negative reactions in the 'Correspondence' section, one from a loyal member of the *Association*, Bessie Robson, who stated that the practical consequence of this innovation was that she "could no longer lend *Le Maître Phonétique* to her students' and the other from his long-time friend and collaborator, Danish scholar Elna Simonsen. She considered that Passy was "assassinating his own work" because, pedagogically, this orthography made no sense.

The study of subsequent issues of the journal reveals that Passy and Jones immediately abandoned the plan to publish all official articles in the new phonetic spelling, returning smoothly to conventional broad transcriptions very much in the Sweet tradition but maintaining now well established IPA symbols. What however remained of all these debates was the insistence on giving privilege to 'Romanic' symbols as Sweet called them along with a concerting effort to make many languages converge in the adoption of similar symbols from this Latin-based set.

6 Consequences and conclusion

The selective portions of history that we have examined in this paper show that the issue of spelling reform and its pedagogical implications has had an over-riding influence on the devising of the IPA. Despite sharp differences in approach, a consistent belief endured throughout most of the twentieth century—and even beyond—in the central importance of certain Latin-based symbols, which had been present from the very inception of the *Association Phonétique Internationale*. Even if Passy and Jones's more radical positions were defeated, a weakened position characterized by systematic use of 'broad' transcriptions (in Sweet's terminology), the retention of blanks, and standard punctuation, remained prevalent, and was implicitly regarded as an ideal form of spelling. If we wish to wrench away phonology and phonetics from techniques aimed at reducing languages to writing (a formulation used by Pike [22] as a definition of phonemics), we must become fully aware of the history of the symbols and the notational conventions we use.

We believe that the centrality of 'Romanic' symbols has had a considerable impact on the description of languages. A salient example of this has been the choice of /r/ for English, French (at the time) and a host of other languages (whether the variety uses e.g. [r], [r̥], [r̥̥], [r̥̥̥], [R], etc., as a main allophone), which can be shown to be highly problematic on general grounds [23]. But this is far from being the only example. Another case is the more or less systematic selection of /e/ and /o/ as default symbols whenever a language has only two mid vowels, a front unrounded one and a back rounded one. This selection may well be linked to the frequent assumption that if there are only two mid back and front vowels they are 'tense' (or mid-high) by default. But when we look at languages which have only vowels of this type (e.g. Spanish), the range of realizations between [e] and [ɛ] is very wide across phonotactic and dialectal contexts (see the contributions in [24]). A final illustration is the choice of /x/ as a symbol for a consonant which is in fact often more appropriately described as a uvular [χ] in many varieties of Arabic (an observation already made in [25]). These selections of Latin-based IPA symbols often give rise to panlectal descriptions giving the illusion of more symmetry than is actually attested, or constraints which require far more investigation, or the positing of historical processes for which the evidence is lacking. To avoid this, we need to make a fuller use of phonetic and psycholinguistic experimentation in addition to corpus-based phonological and phonetic analyses of languages. This is the only way to talk sensibly about phonological systems, their internal structures, their phonetic instantiations, and markedness statements about the languages of the world.

7 References

- [1] MACMAHON, M.K.C.: *The International Phonetic Association: The first 100 years*, *Journal of the International Phonetic Association*, 16, pp. 30–38, 1986.
- [2] SCHWEITZER, C. and C. DODANE: Du son au phonème : histoire des alphabets. In: C. DODANE and C. SCHWEITZER (eds): *Histoire de la description de la parole, de l'introspection à l'instrumentation*, pp. 189–18, Honoré Champion, Paris, 2021.
- [3] SCHWEITZER, C., C. DODANE and J. LAZAR: L'histoire des alphabets phonétiques du XVIIIe siècle jusqu'à l'API. In: *XXXIIe Journées d'Études sur la Parole, 4-8 juin 2018, Aix-en-Provence, France*, p 356-364. DOI: 10.21437/JEP.2018-41, 2018.
- [4] OHALA, J.J.: A brief historical survey of phonetic-phonological feature systems. In: L.-J. BOË and C.-E. VILAIN (eds): *Un siècle de phonétique expérimentale*, Lyon: ENS Éditions, Lyon, 2010. <https://doi.org/10.4000/books.enseditons.31100>.
- [5] LODWICK, F.: *An Essay towards an Universal Alphabet*, *Philosophical Transactions [Transactions of the Philosophical Society]*, no. 182, June 26, pp. 126–137, 1686.
- [6] DE BROSSES, C.: *Traité de la formation mécanique des langues, et des principes*

- physiques de l'étymologie*, 2 vol., Paris, Chez Saillant, Vincent, Desaint, Paris, 1765.
- [7] MONTMIGNON, J.-B.: *Système de prononciation figurée applicable à toutes les langues*, Royez, Paris, 1785.
- [8] DURAND, J. and C. LYCHE: *Paul Passy, un linguiste révolutionnaire. Réforme de l'orthographe, didactique des langues et alphabet phonétique international*, Lambert-Lucas, Limoges, 2024.
- [9] PASSY, P.: *La phonétique et ses applications*, *Supplement to Le Maître Phonétique*, 7(44), 1929.
- [10] PASSY, P.: *Le phonétisme au Congrès phonologique de Stockholm en 1886. Rapport présenté au Ministre de l'instruction publique*, Delagrave/Hachette, Paris, 1887.
- [11] BERGOUNIOUX, G.: La transcription des dialectes gallo-romans. L'exemple de l'alphabet de Rousselot. In L. SCHØSLER and J. HÄRMÄ (eds): *Actes du XXIXe Congrès international de linguistique et de philologie romanes*, vol 1, pp. 201–211, Éditions de Linguistique et de Philologie, Strasbourg, 2021.
- [12] SWEET, H.: *A Handbook of Phonetics, Including a Popular Exposition of the Principles of Spelling Reform*, Clarendon Press, Oxford 1877.
- [13] SWEET, H.: Normalised phonetic spelling, *Le Maître Phonétique* 21, pp. 71–73, 1906.
- [14] PASSY, P.: Anglais: the woman and her pig, *Le Maître Phonétique* 9(2), pp. 46–47, 1894.
- [15] PASSY, P.: Soyons conséquents, *Le Maître Phonétique* 21(12), pp. 127–128, 1906.
- [16] PASSY, P.: Écriture phonétique simplifiée, *Le Maître Phonétique* 22(3/4), pp. 35–37, 1907.
- [17] JONES, D.: Simplified phonetic spelling, *Le Maître Phonétique* 22(3/4), pp. 37–39, 1907.
- [18] PASSY, P. and D. JONES: *The Principles of the International Phonetic Association*. Supplement to *Le Maître Phonétique* 27, pp. 1–40, 1912.
- [19] JONES, D.: Phonetic spelling of English, *Le Maître Phonétique* 28, pp. 113–116, 1913.
- [20] PASSY, P.: Vie nouvelle, *Le Maître Phonétique* 1(38), pp 1–3, 1923.
- [21] PASSY, P.: Orthographe ou transcription?, *Le Maître Phonétique* 1(38), pp. 9–12, 1923.
- [22] PIKE, K.L.: *Phonemics. A Technique for Reducing Languages to Writing*, The University of Michigan Press, Ann Arbor, 1947.
- [23] ANSELME, R., F. PELLEGRINO and D. DEDIU: What's in the r? A review of the usage of the r symbol in the illustrations of the IPA, *Journal of the International Phonetic Association* 53(3), pp. 1–30, 2023.
- [24] GIL, J. and J. LLISTERRI (eds.): *Fonética y fonología descriptivas de la lengua española*, 2 Vol., Georgetown University Press, 2024.
- [25] LADEFOGED, P. and I. MADDIESON: *The Sounds of the World's Languages*, Blackwell, Oxford, 1996.

“PHONETIC SCIENCES IN RETROSPECT” – A WRITTEN INTERVIEW PROJECT

Jürgen Trouvain¹, Pavel Šturm²

¹*Language Science and Technology, Saarland University, Saarbrücken, Germany,*

²*Institute of Phonetics, Charles University, Prague, Czech Republic*
trouvain@lst.uni-saarland.de

Abstract: This paper introduces an ongoing interview project that explores the personal reflections and professional insights of ten senior scholars in the phonetic sciences. Written interviews were employed as a qualitative method to document the evolution of the field through first-hand accounts. The core objective is to develop a broader historical and conceptual understanding of the phonetic sciences by collecting responses to a structured questionnaire organized around ten thematic areas – from research milestones and methodological limitations to interdisciplinary challenges and personal career reflections. A central methodological decision was to conduct written rather than spoken interviews, ensuring consistency across responses and facilitating direct comparison while allowing participants time for considered reflection. The selection of interviewees sought to reflect diversity across subfields, gender, and geography, although the current sample still shows a Northern Hemisphere. Preliminary findings revealed both shared perspectives and notable divergences, including recurring references to figures like Abbé Rousselot and Gunnar Fant. The responses have already uncovered surprising historical anecdotes and insights. As editing continues toward a planned book publication in 2025, the project remains open to expansion, with additional contributions anticipated for the online edition.

1 Introduction

Although ‘interviews’ are not a common genre in scientific literature, at least not in the phonetic sciences, sharing the experience and knowledge of scholars with long and distinguished careers can offer valuable perspectives to members of various scientific communities. In this article we report on methodological issues of an ongoing written interview project on “Phonetic Sciences in Retrospect: Interviews on the Evolution of the Field”. These interviews include responses of (currently) ten colleagues (beyond 65 years of age) who were – and in some cases still are – active in diverse fields of phonetics.

2 Interviews as a special genre of scientific output

An interview can be regarded as a qualitative research method that relies on asking questions to collect data, often performed in social sciences [1] or in ethnographic research [2-4]. In such contexts, the primary aim is to gather empirical data rather than to elicit personal reflections from scholars on their fields of expertise.

When looking more closely to linguistics, and particularly to phonetics, existing interview projects tend to focus either on single personalities recognized as accepted authorities in their fields, e.g. an interview with the sociolinguist William Labov [5], or on oral interviews with several scholars, such as those available at the oral history website of the Acoustical Society of America, e.g. an interview with Katherine Harris, a longtime member of the Haskins Laboratories, published as a transcript by Bell-Berti [6]. Further examples include transcribed oral interviews with the British-American phonetician Peter Ladefoged [7], the Japanese speech

scientist Sadaoki Furui [8] or the British phonetician Michael Ashby on the emergence of phonetics as a field [9]. Further examples of oral history of speech researchers that not only includes transcripts of the interviews but also the corresponding audio files were collected at the Haskins Laboratories [10].

Yet another approach of interviews in linguistics is exemplified by Gururaja et al. [11], who conducted long-form retrospective interviews with more than twenty researchers in the field and delivered a structured summary. Their purpose was to achieve more *general* insights for their field (i.e., natural language processing) for the future by engaging scholars with extensive experience.

For our project, we invited distinguished colleagues to respond to a written questionnaire structured along ten topics, each containing two to three questions (see Section 5). Included are questions about their specific research field and their views on the progress in the field in the last two centuries (highlighting key milestones and scholars, and even proposing candidates for an imaginary Noble Prize in the field). Further topics address their personal career experiences, including reflections on their mentors and comparisons between the academic environment of the past and the present. Respondents were also asked to comment on barriers they encountered (e.g. overlooked or inaccessible publications) and on crucial theoretical concepts. The final topics refer to methodological limitations, interdisciplinarity, and communicating the relevance of their research to laymen.

3 The Interviewees

So far, we have received responses from ten scholars. Their names will be published in the official formats of the project: a book and a dedicated website. The latter offers greater flexibility, allowing the interview collection to be expanded with additional contributions over time. Currently, four more persons have agreed to participate.

The selection of interviewees required a careful balance of various factors. A very important factor for us was that the diversity of the phonetic sciences was represented. Our ten scholars therefore stand for fields and approaches that range from phonology to conversational analysis, from voice and intonation to rhythm studies, from fieldwork to laboratory observation. Admittedly, finding an appropriate term that describes the scientific area we had in mind proved much more difficult than anticipated. We changed from *speech communication research* to *phonetic sciences* after discussions with our first two interviewees who felt uneasy with the former term. In contrast, the latter term, *phonetic sciences*, is broad enough to cover the intended diversity as reflected for instance by scholars with an activity at the International Congress of Phonetic Sciences (ICPhS) [12].

Nowadays, striving for gender balance is often seen as a standard consideration in academic projects. In previous research [13], we showed that, for first authors of conference papers at ICPhS meetings, this was achieved at the congress in 2007. In our current selection, four of the ten interviewees are female.

Another factor that should not be ignored is the geographical distribution of the interviewees. In [13] we demonstrated that ICPhS meetings have shown a strong bias toward researchers from countries in the northern global hemisphere. A similar imbalance is evident in our current sample, with seven out of the ten interviewees based in North America and the United Kingdom, and only three in (Western) Europe. This is a caveat we acknowledge and aim to address in future expansions of the interview project.

4 Comparability of responses

To enable meaningful comparisons between responses, we recognized the need to ask identical or almost identical questions across interviews. It was also important to us that readers should have the opportunity to consult the original answers and not only our interpretations of the

primary data. For this reason, we decided to use a written questionnaire. Although it might be more natural for phoneticians to engage in spoken interviews (be it face-to-face, by phone or online), we opted for the written format to ensure consistency and comparability, being fully aware that some information that might surface in spoken dialogues will be lost during the process.

We developed a fixed questionnaire that provided a stable structure, ensuring that the order and wording of the questions were identical for all interviewees. Although this is a feature of the so-called *structured interview* [1], we did not follow a typically structured interview where questions are so precise that they require clear answers (yes/no). Rather, our questionnaire was designed with elements of *semi-structured* or even *un-structured* interviews, allowing for greater open-endedness and flexibility.

5 The questionnaire

This section contains the original questionnaire that all interviewees were sent.

“As outlined in the email, our written interview project has the objective to explore the history of how the spoken form of language has been studied over time. We intentionally use the term ‘phonetic sciences’ in a rather broad sense to cover various topics and scholars represented in conferences such as ICPhS. Although some scholars might prefer alternative labels (e.g., research on speech communication, spoken language or speech sciences), we feel that the term ‘phonetic sciences’ expresses well our idea of the diversity of the field and its subfields and interfaces to other fields (such as language acquisition, cognition, engineering and many more).

Summary of instructions:

- There are ten sections, with two to three questions in each section.
- Please keep your responses to each section within 600 words (roughly one page).
- Please try to avoid overly complex or formal language in favour of a more engaging style.

5.1 The Field

Let’s begin by exploring your perspective on the field. Specifically, we would like to gain insights into how you would define your area of study within the broad field of the phonetic sciences, how you became interested in it, and your thoughts on what should be included in – and perhaps what excluded from – the study of spoken language.

- How would you define your field of study within the phonetic sciences?
- How did you come to the field?
- What activities and/or types of material are part of the study of spoken language and what sort of activities and/or types of material are NOT part of it?

5.2 The Origins

There is disagreement about when phonetic questions began to be studied scientifically. Firstly, what scholars contributed most, in your view, to the initiation of modern phonetic sciences, probably in the 19th century? Secondly, what were the key moments and ‘firsts’ that shaped the field in its early years (such as important books, influential schools, first attempts of measurement or observation, theoretical breakthroughs, and so on)? And thirdly, are there any crucial precursors that predate the 19th century that you believe are important to consider?

- Which scholars did contribute most to the initiation of modern phonetic sciences in the 19th century in your view?
- What specific moments and other ‘firsts’ come to your mind as important milestones in this era?

- Would you like to suggest examples of crucial precursors in the field before the 19th century?

5.3 Groundbreaking Development

Let's now focus on pivotal events and influential figures that have shaped the field later than those mentioned so far. Could you provide your perspective on the most critical 20th-century milestones that have propelled your area of study forward? Additionally, just envision a Nobel Prize being awarded in your field. Can you share with us who in the 20th century you would have nominated and why (regardless of whether you have personally met that person)? Also, we are interested in your thoughts on the future of the field: Can you think of a future Nobel Prize winner in the 21st century, what achievements and contributions do you believe should be recognized? (There is no need to specify a particular name.)

- According to you, what are the most important (max. three) 20th-century milestones in the phonetic sciences that advanced your field?
- Imagine that a Nobel Prize were awarded in your field: who in the 20th century would you have nominated and for which reason?
- Think of a future Nobel Prize winner in the 21st century in your field: for which reason and achievements should this person or team be nominated?

5.4 Mentors and Unnoticed Neighbours

Here we would like to know who were the mentors or influential voices that helped shape your professional life in the early stages. These could be individuals you've personally collaborated with or authors whose works have had a significant impact on your research. In addition, we are interested in knowing about individuals outside your field in a narrow sense who you believe deserve to be better known in the phonetic sciences. These could be people who have made noteworthy contributions to related fields or whose work has inspired you in some way.

- Who were your mentors that directed your professional life in the early stages?
- What interesting project of theirs inspired you?
- Which people outside the field of your specialization would deserve to be better known in phonetically oriented communities?

5.5 The Time Machine

Let's take a journey in a metaphorical time machine. In the beginning of your career, what are the main differences you have noticed, compared to now, in terms of research questions, methodological aspects, data collection, or other aspects? What possibilities and exciting developments were there then, and how have they evolved over time? Continuing the journey into the future, where do you see the field heading in the next 10 to 15 years? What emerging trends or technologies do you think will be most influential?

- Comparing the beginning and the end of your career, what are the main differences?
- What were the possibilities and exciting developments then, in your early days, and what are they now?
- Where will we be in 10 to 15 years from now?

5.6 Barriers

It is important to acknowledge the potential barriers that can impede progress. One such barrier is the issue of forgotten or inaccessible publications, particularly those due to language barriers or the lack of (free) online availability. Another challenge is the problem of research findings that go unpublished due to a lack of significant effects. From a different perspective, it is also

important to consider the most common errors in publications and conference presentations that can hinder the advancement of the field. In your opinion, what is the most applied error that you see and how can it be avoided?

- Is it a problem for the progression of the field that some publications are forgotten, or inaccessible due to a language barrier or the lack of (free online) availability?
- Is it a problem for the progression of the field that research findings are not published due to a failure to find significant effects?
- What bothers you the most (this can happen on various levels) as a repeatedly observed shortcoming in publications and conference presentations?

5.7 Concepts

Speech, as a form of communication, is undeniably intriguing. It is remarkably fast despite the complex motor control it demands from humans. In addition, speech production and perception seem to occur almost without conscious effort in most interpersonal interactions, conveying both linguistic and non-linguistic information (for example speaker age). From your perspective, which theoretical concepts related to the properties of speech do you consider the most critical in underpinning our understanding of spoken language? These concepts, which can manifest at various levels, shed light on the intricate mechanisms involved in speech production, perception, and cognition.

- What are for you the most important underlying concepts (max. three) observing and explaining speech?
- Which of these concepts do you consider as possibly useful for other disciplines?
- Which of these concepts require more attention in further research and/or in practical applications?

5.8 Scratching the Surface

There are still limitations in the observation of spoken language. For instance, although we know a lot about laboratory speech, how do laboratory findings translate to real-world scenarios? Do we find the same patterns in less-controlled speech outside laboratories? Additionally, the dominance of English in the field poses challenges. While English is the most extensively described language and has many language models based on it, other languages offer a wealth of phenomena that could question the findings based on English. How problematic is this in your view?

- How well do findings from laboratory speech reflect the ‘outer world’? That is, do we find the same patterns in less-controlled speech outside laboratories?
- Languages offer a wealth of phenomena that might contest many views based on the findings from English (or other well-researched languages). How problematic is this English centricity in your view?

5.9 Thinking across Disciplines

Interdisciplinary approaches may have benefits in many scientific disciplines. In this context, we invite you to explore the advantages and challenges of being an interdisciplinary speech scientist, as well as potential methods and scientific disciplines that would be beneficial in the phonetic sciences and could be integrated into the field.

- What are the advantages and what are the challenges of being an interdisciplinary speech scientist in your view?
- Which methods that you see in other fields, from medicine to sociology, would be beneficial in the phonetic sciences?

- To which scientific discipline should we make a closer connection?

5.10 Communicating the Relevance

Our research has the potential to significantly impact society, but it is often difficult to convey its relevance to the public. How can it be addressed? What do you consider to be a good reason to convince laypeople that your research is not only ‘nice to have’ but useful for society as well? Finally, has it been easier to convince others of the relevance of our field in the past?

- What is in your view a good reason to convince laymen that your research is not only ‘nice to have’ but useful for society as well?
- Was it easier to convince others of the relevance of our field in the past (say 50 years ago)?”

6 Editing process

The editing process comprises at least three main steps. First, we established contact with the potential interviewees, and, if willing to participate, we sent them the questionnaire.

Second, once the completed questionnaires were returned, we, as editors, reviewed the text for clarity and formal mistakes. We then returned a typeset version of their responses to the interviewees, along with requests for literature references where needed to enhance consistency across interviews. At this stage, interviewees were also given the opportunity to make final refinements or changes to their responses. We also asked for an informed consent to use their answers in both publications: book and website.

The actual publication represents the third stage. While the interviews may appear consecutively in the book or as an alternative having the ten topics as ten chapters, the website offers a more flexible approach. For instance, individual questions can be selected, allowing users to view and compare responses from multiple interviewees for easier analysis.

7 Selected findings

The editing process is still underway, with final versions expected in summer 2025. As providing coherent and balanced summaries requires a thorough comparative analysis of the responses, we currently present only brief summaries of two selected topics that are easier to report.

The responses to Topic 2 (“The Origins”) reveal a broad consensus around key figures and developments in the early history of phonetic sciences. Abbé Rousselot was the most frequently cited scholar credited with initiating modern phonetic research. As for earlier foundational influences predating the 19th century, Wolfgang von Kempelen was mentioned several times as a crucial precursor. When identifying key moments and ‘firsts’ that shaped the field in its early years, interviewees highlighted three main areas. First, the role of instrumentation, particularly the transformative impact of the spectrograph; second, the development of the International Phonetic Alphabet; and third, the conceptualisation of phonological fundament provided by the German neogrammarians.

The responses to the question about hypothetical Noble Prize nominees in the 20th century revealed a range of prominent figures in the phonetic sciences. Gunnar Fant received the most mentions, cited by three interviewees. Peter Ladefoged, Alvin Liberman, Björn Lindblom and Ken Stevens were each nominated twice. Other individuals mentioned once included Joseph Greenberg, Osamu Fujimura, Eli Fischer-Jørgensen, Mark Liberman and John Ohala.

The Noble Prize question also elicited some unexpected and intriguing responses. One interviewee noted, for example, that “in the early twenties, professor Chaillou of the Institut Pasteur in Paris had proposed Rousselot as a possible Noble prize for physics.” (In fact, historical

records show that Rousselot was nominated in 1914 in the category of “Physiology or Medicine”, although he did not receive the award.)

Another surprising detail came from an interviewee who worked with Jimmy Gene Harris, a trained phonetician and linguist

“who – working with an agency – has actually earned a Nobel Prize. Jimmy Gene Harris was awarded a certificate of the Nobel Peace Prize in 1981 ‘in recognition of devoted service’ to the Office of the United Nations High Commissioner for Refugees, who won the prize for their action in the Cambodian refugee crisis of 1980.”

8 Conclusion

This preliminary report on our ongoing interview project “Phonetic Sciences in Retrospect” reveals two features that set it apart from other interview-based research and, to our knowledge, represent novel contributions. First, the project was designed and conducted using *written* interviews. Typically, interviews involve *oral* interactions, while transcriptions (or selections of it) can be used for printed versions, sometimes alongside the audio recordings. In contrast, our written interview responses are not the product of a spontaneous or semi-spontaneous (when questions are known before) interaction, providing the opportunity to dive into more details or follow up on specific points during the interview process.

The second methodological peculiarity is that our project involves parallel interviews rather than single interviews in isolation. Although each participant was interviewed individually, and the responses can stand alone as separate interviews, the use of identical questions across all interviewees enables direct and meaningful comparisons between their answers.

An interesting by-product of this interview project is the emergence of anecdotal information as those related to the Nobel Prize that are rarely found in academic publications. However, at the current stage, it is difficult to comment on the content in detail. Preparing summaries of this analytical work would require a level of effort that we could not provide before the actual publication. Nevertheless, we anticipate recurring themes, including consensus on pioneering work and reflections on key historical changes. These can serve as orientation lines for future research, as well as a critical lens through which researchers might reassess their own academic histories.

9 Acknowledgements

The authors thank all scholars who participated in this interview project. This publication was supported by the Cooperatio program of Charles University, Linguistics research area, implemented at the Faculty of Arts of Charles University.

References

- [1] KNOTT, E., A. H. RAO, K. SUMMERS, and CH. TEEGER: Interviews in the social sciences. *Nature Reviews Methods Primers* 2, 73, 2022. <https://doi.org/10.1038/s43586-022-00150-6>.
- [2] SPRADLEY, J. P.: *The Ethnographic Interview*. New York: Holt, Rinehart and Winston, 1979.
- [3] DE FINA, A.: *The ethnographic interview*. In: K. Tusting (ed.): *The Routledge Handbook of Linguistic Ethnography*. Routledge, 2019. <https://doi.org/10.4324/9781315675824>
- [4] KIMBER, L. R., and E. DAIRON: *Ethnographic interviews*. In: L. R. Kimber, F. Badache, and L. Maertens (eds.): *International Organizations and Research Methods: An Introduction*, pp. 48–55, University of Michigan Press, 2023.
- [5] GORDON, M. J.: Interview with William Labov. *Journal of English Linguistics* 34(4), pp. 332-351, 2006.

- [6] BELL-BERTI, F.: Interview with Katherine Harris, 2001. <https://www.aip.org/history-programs/niels-bohr-library/oral-histories/33686>
- [7] KAYE, A.S.: An interview with Peter Ladefoged. *Journal of the International Phonetic Association*, 36(2), pp. 137–144, 2006. doi: <https://doi.org/10.1017/S0025100306002519>
- [8] FURUI, S.: An oral history conducted in 2021 by M. A. Hellrigel, IEEE History Center, Piscataway, NJ, USA. Interview #872 for the IEEE History Center, The Institute of Electrical and Electronics Engineers, Inc. https://ethw.org/Oral-History:Sadaoki_Furui
- [9] ASHBY, M., and J. MCELVENNY: The emergence of phonetics as a field. In: J. McElvenny (ed): *Interviews in the history of linguistics: Volume I (History and Philosophy of the Language Sciences 6)*. Berlin: Language Science Press, 2022. doi: 10.5281/zenodo.7092391
- [10] Haskins Laboratories Oral Histories and Transcriptions-bak. Compiled by Carol Fowler and Donald Shankweiler. Transcripts and oral histories from: Alan Cooper, Alice Dourian, Arthur Abramson, Carol Fowler, Craig Cooper, Donald Shankweiler, Kathy Harris, Len Katz, Lloyd Morrisett, Maxine Singer, Michael Studdert-Kennedy, Michael Turvey, Phil Rubin, Cyrus Bacchi, Stuart L. Marcus and Nigel Yarlett. <https://haskinslabs.org/about-us/history/oral-histories-transcriptions-bak>
- [11] GURURAJA, S., A. BERTSCH, C. NA, D. G. WIDDER, and E. STRUBELL: To build our future, we must know our past: Contextualizing paradigm shifts in Natural Language Processing. *Proc. Conference on Empirical Methods in Natural Language Processing*, 16 pp., 2023. <https://aclanthology.org/2023.emnlp-main.822/>
- [12] TROUVAIN, J.: 100 years of diversity in the phonetic sciences – Louise Kaiser, Ilse Lehiste and Anne Cutler. *Proc. 6th Workshop on History of Speech Communication Research*, Budapest, pp. 9-14, 2024. doi: 10.21437/HSCR.2024
- [13] TROUVAIN, J.: On the evolution of ICPhS between 1932 and 2019. *Proc. 20th International Congress of Phonetic Sciences*, Prague, pp. 3236-3240, 2023. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2023/full_papers/856.pdf

**FROM SOCIOPHONETICS TO MUSIC AND LITERATURE:
THE MISSIONS OF JEAN PSICHARI AND HUBERT PERNOT TO CHIOS
(1886-1899)**

Christophe Corbier

CNRS-IReMus UMR 8223, France

christophe.corbier@cnrs.fr

Abstract: The archives of Jean Psichari and Hubert Pernot preserved in Paris and Athens show how the two French linguists anticipated the methods of sociolinguistics and sociophonetics. Their missions to Chios are now known thanks to documents that are in the process of being published: a set of handwritten cards by Psichari (1886) and the sound recordings made by Pernot in 1898-1899. These recordings were found and identified in 2011 in Paris and in 2020 in Geneva. The Psichari files show that Psichari developed a method combining phonetics and ethnographic and social study. Particularly, he gave importance to stigmatisation and phonetic variations according to context. Pernot's investigations show a similar interest in the contexts of Chios. But experimental phonetics applied to the dialects of Chios distanced Pernot from sociolinguistics. Psichari and Pernot's research in Chios gave rise to specific musical and literary productions. For Psichari, phonetics and linguistics formed the basis of a new literary language that would bring science and art together and create a national language. The collection of 114 *Mélodies populaires de l'île de Chio*, by Pernot and Paul Le Flem, was similarly used for artistic and political purposes by French musicologists who wanted to combat the Ottoman Empire and create a new musical language.

Antoine Meillet, Marcel Cohen, André Martinet are generally cited among the French precursors of sociolinguistics and sociophonetics [1], [14], [26]. In contrast, Jean Psichari and Hubert Pernot do not appear in the genealogy of the founders of these disciplines. Nevertheless, I would like to demonstrate how their research, focused on dialectology, was important for Greek phonetics and linguistics, and what consequences it had in the literary and musical fields. This relates to the well-known question of transcription, which brings us back to the canonical problem of the relationship between orality and writing. The subject matter is therefore complex, intersecting phonetics with the history of linguistics, literature, and music in France before 1914.

As we know, Pernot was a pioneer of sound recording, thanks to his missions to Chios in 1898 and 1899, thirty years before his recording campaigns in Bucharest, Prague, Athens, and Paris. In 1928, alongside Béla Bartók and László Lajtha, Pernot participated in the first Congress of Popular Arts and Traditions in Prague, a significant milestone in the development of musical ethnology [3], [10]. Conversely, the importance of Psichari's work on dialects has long been acknowledged in the development of Greek phonetics [12: 179]. As early as 1885-1886, Psichari was examining crucial problems not only for phonetics and linguistics but also

for sociolinguistics and sociophonetics, such as phonetic variations, linguistic changes, and social contexts, as well as the production and reception of speech. Psichari and his pupil Pernot, who studied phonetics with Rousselot, also raised the question of the value of dialects in Greece. They opposed the centralization around the “Katharevousa,” which spread with the expansion of schooling, the press, and administration, as Meillet highlighted as early as 1913 in *Aperçu d'une histoire de la langue grecque* [15: 241-287], [16: 340].

In the 1880s, Jean Psichari's method was based on the works of Gaston Paris and his students, whose articles appeared in the journal *Romania*, as well as on the works of Jules Gilliéron, Abbé Rousselot, and other scholars who published their studies in the *Revue des patois gallo-romans*, founded in 1887. It is easy to understand why, in 1913, Psichari, who had been familiar with these issues for thirty years, wrote to Meillet about Marcel Cohen's book, *Le Parler arabe des Juifs d'Alger* (1912): “I have leafed through this book with delight: it is of general interest, and I shall not fail to quote it in connection with Greek¹.” Moreover, Psichari, Meillet, and Cohen met at the Société de Linguistique de Paris (SLP), which Pernot joined on 26 January 1895 under the patronage of Psichari and Victor Henry, at a meeting chaired by Abbé Rousselot. Between the École Pratique des Hautes Études (EPHE), the Collège de France, and the SLP, the first drafts of a method centered on the study of dialects, phonetic variations in circumscribed geographical areas, interactions between individual speakers and social and/or ethnic groups, and the socio-political significance of this type of research took shape.

1. Jean Psichari's Mission to Chios (September-October 1886)

Before examining Pernot's methodology, we must revisit Jean Psichari's mission in 1886. From 8 August to 29 October 1886, Psichari traveled to the Ottoman Empire. He stayed in two places: Constantinople and Chios. Born in Odessa in 1854, he had never previously lived in Greece or Chios. From 1860 to 1863, he lived with his father in Constantinople. Subsequently, he was raised in Marseille and Paris with his maternal grandmother. An “Agrégé de grammaire” in 1881, he was a pupil of Gaston Paris, Louis Havet, and Arsène Darmesteter at the École Pratique des Hautes Études. He became Ernest Renan's son-in-law in 1883. In April 1885, the young linguist was given a course on the Greek language at the EPHE. His aim was to show the historical development of the Greek language from Antiquity to the modern era, focusing on medieval Greek [12], [19], [20], [21].

A year later, he was asked to represent France at a literary congress organized by the Constantinople Literary Syllogue, scheduled to take place in August 1886. However, the Ottoman authorities canceled the event a few days prior to its opening. Psichari used this time to explore Greek libraries and read the manuscripts needed to compile a historical grammar of Greek language. In September 1886, he traveled to the island of Chios, following in the footsteps of Karl Krumbacher, whom he had praised in 1885 for his book *Beiträge zu der Geschichte der griechischen Sprache* (1884) [27]. Krumbacher visited Chios during a trip to Greece in 1885, and the German scholar, who became a friend of the French linguist, wrote three chapters about his stay on this island in his book *Griechische Reise* (1886) [13].

¹ Letter from Jean Psichari to Antoine Meillet, December 27, 1913, Fonds Antoine Meillet, Collège de France, 83 CDF 6/230.

Psichari spent more than a month in the south of the island, conducting an extensive field survey among the inhabitants. He visited several villages in the south of the island: Pyrgi, Nenita, Olymbi, Mesta. According to his wife Noémi Renan, he worked relentlessly throughout this period [19: 185-187]. The dozens of cards on which he recorded his observations bear witness to the extent of his research. These cards are now in the Psichari Archives (Benakios Library, Athens) and are exceptionally valuable for understanding how Psichari engaged in what can already be seen as a survey of sociophonetics².

In the 154 cards about Pyrgi, Psichari noted the phonetic phenomena he listened to. Most of the speakers were women, whom he heard in everyday situations such as wakes and conversations. He asked them to tell stories and sing songs, noting words and sentences on cards. He often added comments and drew conclusions from what he heard. Psichari found that the women of Pyrgi played a central role in preserving the dialect. Among the women he met, he noticed a desire to protect the ‘patois’ from all the ‘contaminations’ of the standard language.

As early as 1886, he also identified an important factor in understanding the interactions between the standard language and the Pyrgousan ‘patois’: the mockery directed at women who spoke the dialect of their village. Mockery was a crucial factor he had to consider to gain access to the dialect of Pyrgi: the stigmatization of women who spoke ‘choriatika’ (peasant-style) by individuals who had been to Constantinople, Smyrna, or the ‘chora’ (the main town in Chios) posed an obstacle to the search for the dialect in its ‘purest’ form.

“They all tell me that τουέτσιν is not said εν έρκεται καλά [it does not suit well], that those who say it are κουτσόγλωσσοι [crooked tongue] and that they are laughed at (card no. 17).

“The language has allowed itself to be marred by the κοινή [Koine] because of the men, the women on duty at C[onstantino]P[le], the mockery of which they are the object. [...] The women complain about this mockery and say that if they hadn't been mocked, the language would never have changed (card no. 18).

“In many meetings, there are people who make fun of the fact that the language is too peasant (card no. 76).

“Visiting the blind woman. [...] She tells me that when someone speaks κουτσόγλωσσα [crooked tongue] they are laughed at (card no. 90).

On the basis of these observations, Psichari distinguishes between men and women, the former being agents of linguistic transformation while the latter preserve speech in its traditional form. Thus, he was able to formulate the object of his investigation at the very beginning of his research:

“Knowing how to disentangle even under the women's tongues the substance of the dialect, expurgated of the contamination of the κοινή [Koine]. I would often say to them: ah! je vous y prends, vous ne parlez plus pyrgousain and indeed, they would pick themselves up and say to me: πολιτικά είναι” [it comes from Constantinople] (card no. 21).

This research led Psichari to identify two Pyrgousan dialects: an ‘old’ and a ‘new’ one. He proposed the hypothesis of an indefinite evolution of the new Pyrgousan, characterized by

² « Chios », I.I.049.004-006, Psichari Archives, Benakios Library, Athens.

its permanent, daily ‘deformation’ under the influence of the κοινή (Koine) (no. 147). To connect with the women who spoke to him, Psichari attempted to communicate in a manner similar to theirs:

“An old woman to whom I was trying to speak Pyrgousan by saying – end of a sentence – “ὁ κόμμος ὅλος” [the whole crowd] laughed, adding, “Really, he speaks quite Pyrgousan” (card no. 110).”

This relationship, based on trust and reciprocity, enabled Psichari to gather tales with apparent ease:

“Another tale from the same source: “I did recommend that she say βαρεία [heavy] and κουτσόγλωσσα [crooked tongue]” (card no. 72).

The woman who tells the tale is named Irini, daughter of Doullis, and is described as being “14 to 15 years old, was at the χώρα [town]. Poor little girl” (no. 70). He also recorded other details regarding her name, age, gender, level of education, and kinship, demonstrating his sensitivity to the phonetic differences between children and adults.

2. Birth of Greek Sociophonetics in Paris

All this fieldwork forms the basis of the literary and scientific works that Psichari devoted to the question of dialects and Greek between 1887 and 1892. Upon his return to Paris in 1886-1887, he devoted one of his EPHE lectures to the Pyrgousan dialect [28]. Following an paper in the *Revue des patois gallo-romans* in 1888 (“Quelques observations sur la phonétique des patois et leur influence sur les langues communes”), he published a report in the *Nouvelles Archives des missions littéraires et scientifiques* in 1890, and he mentioned his research in Chios in the introduction to *Études de philologie néo-grecque* (1892).

First and foremost, Psichari was able to show the difference between the dialects spoken by the Greek population and the ancient languages (both dialects and koinè), and to verify the validity of several ‘phonetic laws.’ For instance, one he outlined in 1888 states: “Two ancient contiguous deaf explosives cannot exist next to each other; the first becomes spirant.” Here, κτ in ancient Greek becomes χτ in modern Greek, and πτ becomes φτ [31: 194-195]. Moreover, he confirmed the existence of an assimilation (which, according to Psichari, linguist Giorgos Hatzidakis considered impossible) in the Pyrgi dialect: πτ becomes ττ, distinguishing this dialect from modern standard Greek [31: 199-200].

A second result of his research was the construction of a theory of centers from which to analyze the formation of modern Greek and the persistence of dialects [31: 214-219]. The unit of investigation is not the family, as in Rousselot’s dissertation about Cellefrouin, but the village. This serves as a ‘microcosm’ [31: 331-332] in which dialects are spoken by a small community, from which a precise map of Greek dialects can be drawn. In a village, which may become a center under the influence of trade and migratory movements, the dialect and the standard language influence each other. This is why it is very important to carry out field surveys to study phonetic variations in social contexts. Psichari, influenced by Hippolyte Taine’s theory of the “milieu” (environment), refers to the ‘psychological environments’ that affect the pronunciation of speakers from Pyrgi [29: XXXV]. He stressed the ‘state of inattention’ of the speakers, which is a crucial condition for phonetic experimentation. According to him, recording the speech of speakers outside their environment—such as in large cities—is futile, as they are subject to the influence of the standard language [29: XXXV].

3. Far from Paris and Psichari: Pernot in Chios

After 1886, Psichari planned to write a summary of his observations about Chios. However, he did not carry out this project and did not have the opportunity to return to the island in the following decades. In 1897, his pupil Hubert Pernot decided to devote his doctoral thesis to the dialects of Chios [9:59], [4:56]: he wanted to make recordings with a phonograph in the south of Chios, in the places where Psichari had conducted his survey twelve years earlier.

Since 1891, Pernot had been considering studying Greek dialects (particularly Tzaconian, which he had studied in the Peloponnese in 1893). He had also trained in experimental phonetics with Rousselot and, in 1891, he intended to publish a study on the dialect of Froideconche, the village where he was born, on the borders of the Vosges and Haute-Saône, in the *Revue des patois gallo-romans* [9:60]: as we see, Abbé Rousselot's dissertation on the Cellefrouin dialect (which served as a model for Louis Gauchat in 1905 [14], [26]) cannot be underestimated in the works of Pernot on Greek dialects [32].

Pernot's method differed fundamentally from Psichari's. According to Psichari, the machine had a profound influence on phonetics: it eliminated the 'state of inattention' and forced the speaker to concentrate on the production of his speech [29: XXXIV]. The 'speaking subject' pays attention to what he says and tends to adopt the standard language by erasing the pronunciation of his environment. Psichari believed that the phenomena he heard could not have been recorded with a phonograph or any other instrument. On the contrary, Pernot first made recordings of the speakers with a phonograph (a Columbia graphophone weighing 9 kg [22], [4: 57]). Then, he studied them in Paris thanks to 'samples of dialects' [23: 4]. "It's quite different from making people talk in the air and writing down what you hear in your notebook," he explained to his mother in March 1898³.

The missions to Chios were therefore essential to his experimental protocol. They were both organized at the same time of year, during the summer and autumn of 1898 and 1899. Annie Goffre described how they were prepared and carried out [9: 58-59]. Pernot left Paris for Athens at the beginning of July 1898. After spending two weeks in Athens with his mother and sister Marthe, he arrived in Smyrna on July 22. On the 23rd, he was in Chios, and on the 27th, he reached Pyrgi. Due to the harassment caused by the Ottoman administration, which Pernot described at length in *En pays turc. L'île de Chio* (1903), the recordings did not begin until September.

At the beginning of September, Marthe Pernot brought the phonograph and joined her brother and a young American linguist, William Milwitzky (1876-1956), who was working on Judeo-Spanish language. Field recordings were made until mid-October, at which point Pernot returned to Athens. Then he returned to France at the end of October 1898. In January 1899, he sent his first report to the Ministry of Education, requesting further funding for a second mission to Chios. This mission was scheduled to take place from July to October 1899. As in the previous year, Pernot left for Greece at the end of June 1899, stayed in Athens for two weeks, and arrived in Chios after July 15. This second mission did not apparently encounter the same difficulties with the Ottoman authorities.

³ Letter from Hubert Pernot to his mother, March, 18, 1898, BnF, Département Son Vidéo Multimédia.

In 1898, Pernot went only to Pyrgi and to the villages located in the south of the island. In 1899, he recorded the inhabitants of the north of the island before returning to Pyrgi at the end of his journey. In his account, Pernot gives no precise indication of the chronology of his investigations. Everything suggests that it was only during this second mission that he worked with Konstantinos Kanellakis (1840-1916). Kanellakis was the author of a compendium of the island's traditions, *Chiaka Analecta* (1890), in which he published numerous songs without musical transcription. He had been in contact with Psichari since 1886-1887 [33: 168-169]. Krumbacher painted a glowing portrait of him in 1886 [13: 220-221]. In other words, Kanellakis had extensive experience in fieldwork as well as relations with European scholars.

His role of intermediary between the population of the island's villages and the French linguist is important for two reasons: firstly, because it influenced the speakers, their performance, and their relationship with Pernot; secondly, because Kanellakis was a supporter of the attachment of Chios to the Greek nation-state, which implied a rejection of the Turkish presence on the island. Pernot himself was constantly critical of the Turks, and his run-ins with a gendarme can be partly explained by his hostility. Like Psichari and Kanellakis, he made no secret of his support for the island's reunion with Greece. As a result, in the recordings, Pernot obviously only focused on the Greeks.

4. From Recording to Analysis

The cylinders kept at the Bibliothèque nationale de France and the Musée d'Ethnographie de Genève need to be analyzed today. Firstly, 92 of the 146 cylinders recorded in 1898 and 1899 were found in 2011 and 2020. We now have two-thirds of the cylinders. It is possible that the last third has been lost forever, as it may correspond to all the cylinders that were accidentally destroyed between 1931 and 1935.

There are 55 cylinders at the BnF and 37 in Geneva. The first set contains stories, a conversation, and a 'phonographed letter'; recordings of Greek demotic writers Alexandre Pallis, Ioannis Polemis, and probably Georges Souris; the voice of 'Madame Legrand,' probably Emile Legrand's wife; a voice indicated as that of 'Madame Psichari'; popular songs from Greece and France; and test cylinders. The second set contains a number of dance tunes performed on the violin, songs, and short poems [4]. The cylinders containing the tales and conversations often show more wear and tear than the cylinders containing the songs.

Each cylinder is placed in a blue or brown box. The colors suggest the date: cylinders recorded in 1898 are placed in a brown box, while 1899 cylinders are in a blue box. Round labels are affixed to the top and base of the boxes. These labels contain information written in either black ink or black pencil, probably at two different times. Most of the information corresponds to the audio content of the cylinders. In classifying the recordings, Pernot followed Psichari's recommendations. He wrote on the labels the identity, age, and village of the speaker, the first words of the songs or stories, or the type of dance performed by the violinist. When the identity was unknown, he indicated the gender.

Experimental phonetics applied to the dialects of Chios constitute the great originality of Pernot's work. Rousselot's instruments provided him with a guarantee of scientific accuracy. The doctoral thesis he defended in 1907 (*Phonétique des parlers de Chios*) stands out from everything that had been produced in previous years on Greek dialects, such as the books by Paul Kretschmer on the dialect of Lesbos (1905) and Albert Thumb on the dialect of Amorgos (1892).

On the other hand, everything that Psichari had noted about environments and contexts tended to disappear. Pernot went in another direction, abandoning the prospect of sociological phonetics outlined by Psichari. This is particularly evident in the analysis of mockery. Pernot notes that the village of Lithi is the target of mockery by the inhabitants of the neighboring villages [23: 72-73]. However, he concentrates on analyzing accentuation in the Lithi dialect and confines himself to the phonetic level: he uses the graphic method of Marey and Rousselot, without entering into a sociological analysis.

Moreover, he made several recordings in Paris. As he himself explained, he brought three people to the Collège de France to work in Rousselot's laboratory: the archimandrite of the church of Saint Stéphane in Paris, Germanos Vassilakis (1859-1929), from Alatsata (then a Greek village on the Eritrea/Karaburun peninsula), and two men, Nestor and Jean Valiadis, from the village of Vessa. They read a tale by Drossinis while Pernot recorded them with the phonograph in order to study the "nature of the accent" [23: 26-27]. He measured the pitches and durations of the words, then transcribed them using Rousselot's phonetic alphabet and a five-line staff.

This type of notation was recommended by Rousselot in his *Principes de phonétique expérimentale*. The aim is to represent the musical accent of dialects using the Western musical system. Pitches and durations are calculated by using tuning forks to highlight intervals and the melodic line of speech. According to Pernot and Rousselot, it would suffice to convert the number of vibrations recorded by the instruments into musical notes to represent the prosody.

What Pernot thought he discovered in those graphs and scores was the existence of a melodic accentuation. While linguists have shown that melodic accentuation has totally disappeared since the second century AD, he believed that a few traces of it remained in the dialects of Chios. In 1907, he devoted several pages to this subject when he studied the relationship between intensity accent and melodic accent. According to him, the graphic method shows that in the dialects of Chios, the initial intensity is combined with a melodic accent. The "score" of a tale also shows that the highest notes carry the accent of the word [23: 54-63; 24: 4-35].

Moreover, in 1946, Pernot asserted that the scansion of poems spoken by Chios speakers would show the survival of a quantitative metric. For example, "political verse" is a "direct descendant of the ancient Trochaic catalectic tetrameter" [24: 36-37]. Pernot notes also two dipodies in a verse heard in Chios. In this quantitative metric which would be inherited from ancient Greek poetry, verses are built on durations and the word's intensity accent plays no role. Finally, dialectology and phonetics confirm the island's Hellenic identity. This brings us to the political and ideological uses of Pernot's recordings.

5. Musical Language and Literary Language

In 1903, Hubert Pernot and Paul Le Flem published *Mélodies populaires de l'île de Chio*. This collection contains transcriptions of 114 melodies. Very quickly, these melodies were included in lectures on music history. French musicologists aimed to explain the harmonic and rhythmic systems of "Greek popular music" and to fight against Turkish influence. In 1904, after a mission to the Armenian territories of the Russian Empire, Pierre Aubry (1874-1910) delivered a lecture about the music of the oppressed minorities in the Ottoman Empire, specifically the Armenians and Greeks. During this lecture, which took place at the École des Hautes Études Sociales in Paris, the audience heard two melodies from the Pernot-Le Flem collection, harmonized by Maurice Ravel at the request of Michel-Dimitri Calvocoressi (1877-1944). In 1906 and 1907, Calvocoressi himself, whose family hailed from Chios, pre-

sented lectures on popular Greek music featuring melodies by Pernot harmonized by Ravel [25; 2: 5].

The Hellenic identity of Chios was clearly emphasized in 1908 by musicologist and composer Maurice Emmanuel (1862-1938). A friend of Pernot and a pupil of Louis Havet and Gaston Paris, Emmanuel was the teacher of Psichari's daughter, Corrie, who would later become his pupil at the Paris Conservatoire after 1911 [7: 250]. In May 1908, Emmanuel gave a lecture on "Greek Popular Music" as part of the "Hellenismos" Association, founded in Paris in 1904 to defend Greece against Turkish and Bulgarian influences. The audience included Théodore Reinach, Théophile Homolle, and the Greek ambassador to France, Nikolaos Deliyannis, as well as Jean Psichari, Hubert Pernot, and Archimandrite Germanos Vassilakis. Emmanuel's lecture was published on June 1, 1908, in the association's journal, *L'Hellénisme* [6].

On June 10, 1908, Pernot congratulated Emmanuel⁴. Vassilakis and Psichari must also have been pleased. Drawing on his expertise as a theorist and historian of Greek music, Emmanuel integrated the melodies of Chios into the framework of Hellenism. His theory is based on the idea of the continuity of music and language since Antiquity. Unaware of the situation on the island, Emmanuel never visited Greece. In his lecture, he disregards the context in which the melodies are performed and shows little interest in phonetics. Emmanuel was not only a musicologist; he was also a composer who sought to regenerate Western music through contact with Greek music, which he viewed as a coherent and homogeneous system since Antiquity. According to him, modern composers must leverage the history of musical language if they intend to create works rooted in the Greco-European tradition [2:329-385]. This presupposes knowledge of ancient Greek modes and rhythmic patterns, which are perceptible in the musical language used by the musicians and speakers of Chios.

This scientific and musical project parallels the linguistic regeneration project formulated by Psichari in the 1880s and 1890s. Psichari and Emmanuel share much in common. Not only were they both students of Louis Havet and Gaston Paris, but for both the writer and the musician, linguistics and creation cannot be separated. Just as Emmanuel rejected academic music, Psichari rejected the artificial language of the purists, who were not able to communicate effectively with the people. Greek linguistics was employed to combat the "purified" scholarly language. What Psichari aimed to achieve with his travelogue, *To Taxidi mou* (1888), was a linguistic reform based on applying the phonetics to literary writing. For him, Greece needed a national language, which could only be the language spoken by the people [18]. The demotic language should not merely serve as a means of communication; it had to guarantee the authenticity of the national literary language. According to Psichari, "modern Greece, strictly speaking, does not yet have its own literature" [30: XIII]; so the writer's duty is to demonstrate the evolution of Greek through scientific means and to create a literary language based on phonetics and linguistics. Ultimately, "science and literature merge" [30: XIII].

⁴ Letter from Hubert Pernot to Maurice Emmanuel, June, 10 1908, Antony, Maurice Emmanuel Archives.

Psichari's phonetic observations in 1886 could have steered the study of dialects and their particularities towards sociophonetics and sociolinguistics. Similarly, Pernot's sound recordings could have paved the way for a musical ethnology focused on sound and speech. But the tropism of the written word was more powerful than listening, both for Psichari, who used phonetics to attack purists and create a new literary language, and for Pernot, who remained committed to transcription and experimental phonetics. It was not until the second half of the century that sociolinguistics and ethnomusicology developed, as technological advancements combined with the influence of the social sciences opened up new perspectives for phonetics. Nevertheless, we can see today from the Psichari and Pernot archives that their three missions to Chios between 1886 and 1899 were clearly key moments in the development of sociolinguistics and sociophonetics. In particular, the Pernot archives, recently rediscovered and soon available online, will open up new perspectives on the work of French linguists at the end of the 19th century.

References

- [1] CANDEA, M. and C. TRIMAILLE: *Phonétique, sociolinguistique, sociophonétique: histoires parallèles et croisements*, *Langage & Société*, n° 151, pp. 7-25, Mars 2015.
- [2] CORBIER, C.: *Poésie, musique et danse. Maurice Emmanuel et l'hellénisme*, Classiques Garnier, Paris, 2010.
- [3] CORBIER, C.: *The Distant Past: Greek music between archaeology and ethnology (1926-1936)*. In: GUILLEBAUD C., S. EMERIT, and J. JUGAND (ed.): *Singing the Past, Music and the Politics of Memory*, pp. 93-114, Nanterre, Presses Universitaires de Paris Nanterre, 2023.
- [4] CORBIER, C., and KOKKONIS, G.: *Du phonographe au musée virtuel: les archives sonores d'Hubert Pernot, hier et aujourd'hui*, *Revue de la BNU de Strasbourg*, pp. 53-59, Mai 2025.
- [5] CORNEJO, M., and DIAMANTOPOULOU-CORNEJO, D.: *Quelques précisions sur la genèse des Mélodies populaires grecques harmonisées par Maurice Ravel: des dix mélodies initiales aux six mélodies éditées*, *Cahiers Maurice Ravel*, pp. 149-192, 2010.
- [6] EMMANUEL, M.: *La musique populaire grecque*, *L'Hellénisme*, pp. 1-3, Juin 1908.
- [7] EMMANUEL, M.: *Lettres choisies 1880-1938*, Vrin, Paris, 2017.
- [8] GOFFRE, A.: *Hubert Pernot et la chanson populaire grecque*, *Connaissance hellénique*, pp. 44-52, Janvier 2004.
- [9] GOFFRE, A.: *Hubert Pernot et la chanson populaire grecque*, *Connaissance hellénique*, pp. 58-63, Avril 2004.
- [10] GOFFRE, A.: *Hubert Pernot et la chanson populaire grecque*, *Connaissance hellénique*, pp. 22-29, Juillet 2004.
- [11] IGLESIAS, N.: *Sociolinguistique historique*, *Langage & Société*, pp. 315-318, 2021, <https://doi.org/10.3917/l.s.hs01.0316>.
- [12] KRIARAS, E.: *Psycharis: Idees, agones; o anthropos*, Estia, Athènes 1981.
- [13] KRUMBACHER, K.: *Griechische Reise*, Hettler, Berlin, 1886.

- [14] LABOV, W.: *The Social Stratification of English in New York City*, 2nd edition, Cambridge University Press, New York, 2006.
- [15] MACKRIDGE, P.: *Language and National Identity in Greece 1766-1976*, Oxford University Press, Oxford, 2009.
- [16] MEILLET, A.: *Aperçu d'une histoire de la langue grecque* [1913], 7^e édition, Klincksieck, Paris, 2004.
- [17] MITSOU, M.: *Epistoles tou Psichari ston Karl Krumbacher*. In: I. FARINOU-MALAMATARI (ed.), *Psycharis kai i epokhi tou*, pp. 459-473, Thessaloniki, INS, 2005.
- [18] MITSOU, M.: *Titu Maiorescu, Jean Psichari et la conscience littéraire nationale, Cahiers balkaniques*, 42, 2014, <http://ceb.revues.org/4943>
- [19] ORGEOLET, D.: *Évocation de Jean Psichari vivant, Bulletin de l'Association Guillaume Budé*, pp. 180-197, 1978. https://www.persee.fr/doc/bude_0004-5527_1978_num_1_2_3429
- [20] PAPADOPOULOU, D. P. *Jean Psichari, le linguiste du grec moderne: une carrière française* ». In: M. GIRARD, and C. BÉCHU (ed.) : *La France et la Grèce au XX^e siècle: des archives à l'histoire*, École française d'Athènes, Athènes, 2021. <https://doi.org/10.4000/books.efa.13570>
- [21] PERRAKY, V.: *Jean Psichari à la Société Linguistique de Paris et à l'EPHE*. In: C. TRAUTMANN (ed.), *De la philologie allemande à l'anthropologie française. Les sciences humaines à l'EPHE (1868-1945)*, pp. 185-208, Champion, Paris, 2017.
- [22] PERNOT, H., *Mélodies populaires grecques de l'île de Chio*, Leroux, Paris, 1903.
- [23] PERNOT, H.: *Études de linguistique néo-hellénique I: Phonétique des parlers de Chio*, Chez l'auteur, Paris, 1907.
- [24] PERNOT, H.: *Études de linguistique néo-hellénique II: Morphologie des parlers de Chio*, Les Belles Lettres, Paris, 1946.
- [25] PERNOT, H.: *Études de linguistique néo-hellénique III: Textes et lexicologie des parlers de Chio*, Les Belles Lettres, Paris, 1946.
- [26] PRESTON, D. R. and NIEDZIELSKI, N. (ed.): *A Reader in Sociophonetics*, De Gruyter Mouton, New York, 2010.
- [27] PSICHARI, J.: *Langue néo-grecque. Rapport sur l'École pratique des hautes études, 1884-1885*, pp. 167-169, Paris, 1885.
- [28] PSICHARI, J.: *Langue néo-grecque. Rapport sur l'École pratique des hautes études, 1886-1887*, pp. 195-197, Paris, 1886.
- [29] PSICHARI, J.: *Études de philologie néo-grecque: recherches sur le développement historique du grec*, Bouillon, Paris, 1892.
- [30] PSICHARI, J.: *Autour de la Grèce*, Calmann Lévy, Paris, 1897.
- [31] PSICHARI, J.: *Quelques travaux de linguistique, de philologie et de littérature helléniques 1884-1928*, Société d'édition des Belles-Lettres, Paris, 1930.

[32] RIALLAND, A., C. FOUGERON, and P. ŠTURM: *La succession de Rousselot à travers la correspondance de Pernot et Millet avec Chlumský*, *Transversalités*, n° 171, pp. 37-50, 2024. <https://doi.org/10.3917/trans.171.0037>.

[33] ZAHARI-LOUTRARI, A.: *Konstantinos Kanellakis*, vol. 1, Alpha Pi, Chios, 2017.

FROM THEODORE ROSSET TO FRANK C. CHALFANT: A TRANSATLANTIC BRIDGE FOR THE TEACHING OF PRONUNCIATION

Laura Abou Haidar¹

¹Univ. Grenoble Alpes, Laboratoire LIDILEM, France
Laura.Abou-Haidar@univ-grenoble-alpes.fr

Abstract: This contribution focuses on a very specific episode in the history of language teaching methodologies: the influence of Théodore Rosset and his teaching of pronunciation for foreigners based upon experimental phonetics and an innovative dispositive, and two important elements in how knowledge flowed across the Atlantic. First of all, the presence in Grenoble in 1909 of a very special student, Frank C. Chalfant, who imported to the United States Rosset's teaching methods. Secondly, the fact that soldiers and officers have been trained in Grenoble in the early years of the 20th century. This study sheds a new light on the context that preceded the emergence of the American Speech Training Program (ASTP) in the United States.

1 The role of phonetics in enhancing the attractiveness of teaching French as a foreign language

1.1 The *Comité de patronage des étudiants étrangers* of Grenoble

Modern language teaching in France during the second half of the 19th century was far from being structured and based on common, robust pedagogical foundations: As Galazzi notes, “There was total disagreement as to the purpose of this teaching and the methods to be used. This was due as much to the absence of any tradition as to the recruitment of teaching staff” ([9]:28). However, in the last years of the nineteenth century, a major shift was about to take place, marked by pedagogical innovations in the teaching of French to foreigners.

In 1896, a group of prominent figures and academics from Grenoble in France—deeply concerned with the city's attractiveness and international influence, as well as of the Alps—founded an association with the goal of offering language courses for foreign students. These intellectuals aimed to capitalize on the scientific excellence of the University of Grenoble, which was highly regarded in disciplines such as Literature, Law, and Physics. The association was named *Comité de patronage des étudiants étrangers* (Committee for the Support of Foreign Students, hereafter referred to as CPEE or Committee).

These notables and scholars believed that the promotion and dissemination of French language and culture would be more effective if carried out by former students—who had studied in Grenoble. Their ambition was also to attract a wealthy, international clientele, interested in tourism and alpinism, capable of boosting local commerce and forging global connection.

French language instruction for foreign students officially began at the University of Grenoble in 1896, driven by academics committed to a scientific and modern approach to language teaching. The first student to enroll in these courses was German. It was soon decided to establish two types of programs: annual French courses, primarily for students already enrolled in various faculties of the University of Grenoble, and summer courses, open to an international audience and designed specifically for students attending classes during the four-month summer period, from June until September.

This communication focuses on the summer course audience. Enrollment grew rapidly: in the 1904-1905 academic year, 265 students were registered in the annual courses, while 370¹ attended the summer session (*Comité de Patronage des Etudiants Etrangers*, Annual Report, 1905, pp. 16-17).

Over this period, participants from approximately some forty countries are represented. The students can be broadly divided into three main categories.

- **The German public:** from the very beginning and until the eve of the First World War, Germans formed the largest and most consistent group of attendees at the CPEE's annual summer courses (Figure 1). Historian René Favier, who has studied the CPEE archives—particularly the strong presence of German students at the University of Grenoble during this period—offers some valuable insights to this trend: “The agreement signed with the German government in 1901, which validated the semester spent in Grenoble as part of the law curriculum, suddenly swelled the number of students” ([7]: 97).

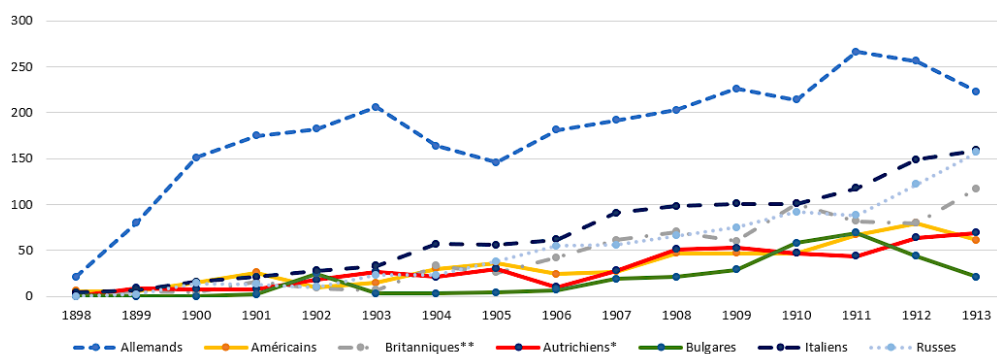


Figure 1. Membership trends of the seven most represented countries (out of 42) in Grenoble CPEE summer courses between 1898 and 1913 (source: Abou Haidar, 2024)

- **Other European and American students:** although significantly less numerous than the Germans, students from several other countries were increasingly represented in the Committee's summer course enrolments. These included a number of European nationalities—Italians, Russians, British, Austrians and Bulgarians—as well as Americans, who were a key target audience for the Committee.
- **Diverse international participants:** the final category includes students from 32 countries, each represented by at least one attendee, with up to twenty students from some of them over the period: Egypt, Japan, Paraguay, Indoustan, Turkey, Chile, Brazil, Australia, Armenia, Argentina, among others.

1.2 Choosing modernity to increase attractiveness: discovering phonetics

The founders of the CPEE were deeply concerned with enhancing the appeal of their courses, believing that their attractiveness depended largely on offering scientific and modern content. As early as the 1902-1903 academic year, a course in modern French phonetics was offered for foreign students. The following year, 1903-1904, the CPEE made significant progress in this area by inviting Paul Passy—an eminent figure in the field of phonetics—to contribute to the development of phonetics instruction.

To make our teaching more complete and to give it greater value, we called upon (...) Mr. Paul Passy, sub-director of the Ecole des Hautes Etudes, one of the creators and one of the recognized masters of the teaching of phonetics. (...) After listening to his

¹ All the historical elements that follow are based on the *Annual Reports of the Comité de Patronage des Etudiants Etrangers* [Archives, Université de Grenoble Alpes], sent to the Recteur d'Académie, in which the teaching, promotion and partnership activities implemented by the CPEE are listed in great detail.

clear and learned words, our listeners took away with them the memory that the teaching of phonetics had its masters in France as well as in Germany, and that the University of Grenoble spared no effort to keep up with the needs of modern teaching.

(CPEE Annual Report, René Reymond, 1904, p. 5)

In 1904, the CPEE appointed a new phonetics instructor, Théodore Rosset, a former student of Ferdinand Brunot, one of the pioneers of phonetic science. Shortly after his arrival, Rosset founded the Grenoble Phonetics Institute. He specialized in teaching French pronunciation for foreign students, employing both the phonograph and the dictaphone a part of his instruction. In 1905, he began formalizing his method in a work titled *Exercices pratiques d'articulation et de diction*.

The Committee closely followed developments in pronunciation pedagogy and recognized the innovative nature of Rosset's approach. It praised both the individualized method of instruction and the use of technological instruments, as highlighted in the CPEE Annual Report of 1905:

Efforts were concentrated on teaching pronunciation: every week, two hours were devoted to the theoretical and experimental study of French phonetics; once a week, students were brought together in the afternoon in the laboratory, where they were made to observe and verify the facts and principles expounded in the theoretical course. Practical exercises were designed in a new spirit: reading lessons were individual and methodical. Before moving on to the reading of extended texts, we made sure that each student pronounced all the French articulations correctly, first one by one, then grouped together; if necessary, errors were rectified; once correct articulation of all the sounds had been acquired, we then moved on to the study of diction, and we drew great benefits from the regular use of phonographic auditions. (...) The listeners who took part in the practical exercises had free access every evening to the phonetics laboratory, where they could see for themselves the reality of the errors found in their pronunciation, and where these errors were occasionally corrected with the help of devices.

(CPEE Annual Report, René Reymond, 1905, pp. 5-6)

Rosset's teaching methods, along with their underlying theoretical principles, have been analyzed by [8, 9] and [2]: these scholars examine Rosset's *Exercices pratiques d'articulation et de diction*, whose successive editions—from the first in 1905 to the final edition in 1925—reflect not only Rosset's pedagogical choices for teaching pronunciation as part of French language instruction, but also broader scientific developments that were actively debated among specialists in the early 20th century.

In these *Exercices*, the pedagogical approach is described as highly “innovative” [9]: “It took full advantage of the prodigious progress that experimental phonetics had just made with Abbé Rousselot” (ibid., 81). Rosset's method was strikingly modern for its time. Phonetics was conceived as an essential component of language instruction, and the teaching approach was already multi-sensory. Rosset integrated visual and auditory modalities, linking production and perception. His pedagogical method involved:

- Seeing, through the use of devices available in the laboratory for visualizing and analyzing articulatory movements;
- Listening, via the phonograph, employed in both group and individual activities. Rosset conducted daily individual sessions with each student, offering tailored pronunciation practice.

This approach is also prescriptive and rigorous: learners' errors were never overlooked, and students were expected to become consciously aware of their mistakes in order to correct them. At the same time, Rosset acknowledged the variability and subjective nature of auditory perception. In the introduction to the Thesis he defended at the *Faculté des Lettres* of Paris in

1911, he observed: “*We hear what we want to hear*” ([15]:10). This recognition of perceptual subjectivity further justified the use of phonetic instruments (Figure 2) in the classroom. These tools, seen as objective and scientific, were intended to reduce the influence of individual cognitive bias and ensure more accurate pronunciation training.



Figure 2. Théodore Rosset surrounded by his instruments (Coriandre Vilain and Nathalie Vallée, GIPSA-Lab archives, Université Grenoble Alpes, used with permission).

The use of the phonograph, dictaphone and other speech-analysis instruments—combined with listening and rehearsal activities in both individual and group settings—proved particularly effective and appealing to the growing number of learners attending the CPEE in the early 20th century. Phonetics and pronunciation teaching were given significant prominence in the CPEE’s annual reports. This is evident in the detailed attention to pedagogical methods, the strategic recruitment of Théodore Rosset, and the numerous lectures delivered by pioneers of phonetic science, including Abbé Rousselot (1900), Paul Passy (1904), and Ferdinand Brunot (1910).

In the 1907-1908 academic year, Grenoble established a national *Diplôme d’études supérieures de phonétique* (Higher Diploma in Phonetics). According to the CPEE Annual Report (1908: 13), instructors awarded this diploma were expected to become “familiar with the various methods used in phonetic studies and (...) capable of teaching pronunciation”. Ferdinand Brunot served as chairman of the first examination session for this new diploma in 1909.

2 American students: a prime target and a gateway across the Atlantic

2.1 Americans in Grenoble

Since 1896, the CPEE has actively cultivated relationships with American academics from prestigious institutions. In 1896-1897, two professors—Mr. Bishof from Yale and Mr. Alan Marquand, Professor of Art History at Princeton—were invited for a one-month stays. By the early 20th century, the University of Grenoble had developed strong ties with several American universities, most notably Harvard.

From the very beginning of French language instruction for foreigners, American students were well represented: 8 out of 57 students in 1897-1898, and 31 in 1900-1901. Americans were a key target audience for the CPEE. However, the Committee was well aware that Grenoble’s location—far from Paris, where most foreign students first arrived—posed a challenge. To address this, the Committee lobbied the P.L.M. (Paris-Lyon-Méditerranée) railway company to offer discounted fares for American students disembarking on the Normandy coast, and facing a long and costly journey to Grenoble. As early as 1901-1902 (CPEE Annual Report, pp. 20-21), the Committee began negotiations with the P.L.M. company to secure reduced rates for all foreign students, though this discount was not implemented until 1910-1911.

The American students who attended CPEE courses were generally well-educated and professionally engaged. Many of those enrolled in the summer sessions were teachers—university professors, primary school teachers, instructors, inspectors. A number of prestigious institutions were represented, including Yale, Berkeley, Catholic University of America, Cincinnati, Harvard, New York, Boston, Michigan, Cambridge, Colombia, Syracuse... In addition to the civilian participants, several high-ranking military officers also took part, as will be discussed later.

2.2 Frank C. Chalfant: following in Rosset's footsteps, from Grenoble to Washington

In the summer of 1909, Frank C. Chalfant, an American teacher, attended summer courses in Grenoble. Two years later, Chalfant would establish the first language laboratory at an American university, Washington State College.

The archives of Washington State University² contain a manuscript written by C. Chalfant (Figure 3A), likely as part of his application to this institution. Chalfant's academic journey in Europe included visits to several prestigious higher education establishments: Post-graduation in German and French at the University of Marburg in Germany in 1905, the University of Berlin from 1905 to 1907, the University of Reims (spelled Rheims at the time) in the summer of 1908, and finally, the University of Grenoble in June 1909 (Figure 3B).

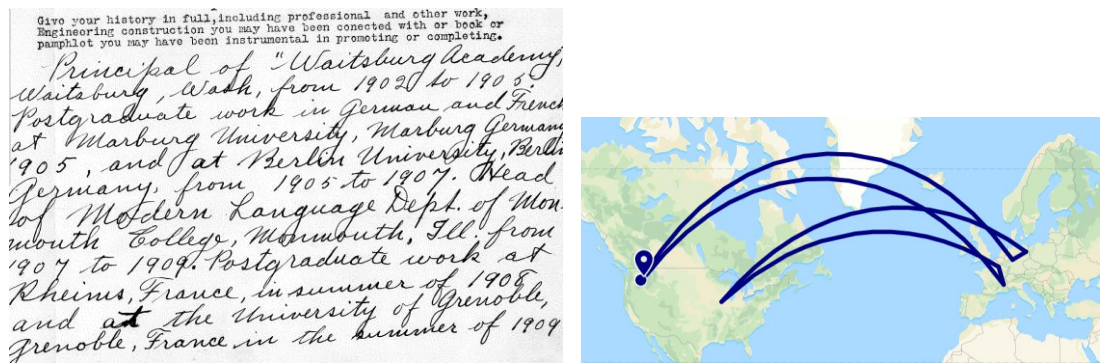


Figure 3. A. Chalfant's manuscript detailing his career prior to joining Washington State College³. B. Chalfant's transatlantic itineraries, 1902-1911 (<https://travelmap.net>)

Chalfant attended the summer courses in Grenoble from June to September 1909. Like all students, he participated in Rosset's classes and was trained in modern techniques for teaching French pronunciation. According to the website of the *Language Learning Resource Center* of Washington State University, where Chalfant became a professor, he discovered the language laboratory in Grenoble. The site notes that Chalfant “*appears to have been the one who brought the idea back to this country*”⁴. Upon his return to the United States, Chalfant established the first university language laboratory at *Washington State College* in 1911-1912 academic year, as documented on the WSC website.

² Acknowledgments to Mr. Mark O'English, archivist, for providing information and documents about Frank C. Chalfant.

³ Manuscripts, Archives and Special Collections, Washington State University Libraries: WSU Human Resources Faculty Personnel Files (UA1987-26, box 3).

⁴ Source: <https://slcr.wsu.edu/llrc-history/>.



53

Figure 4. Illustration published in the 1916 edition of Washington State College's *Chinook* newspaper (Source: <http://content.libraries.wsu.edu/cdm/ref/collection/chinook/id/10821>)

An entire page of *The Chinook*, the Washington State College yearbook, is dedicated to the newly established language laboratory under the title “Snapped in the language laboratory”, featuring a series of photographs taken inside the lab (Figure 4):

- One picture shows a young woman recording herself with a dictaphone using a microphone. This is Ina Wright-Herbst, an “instructor in voice”, who was also a trained singer known for performing at prestigious ceremonies. She participated in both artistic and “educational recitals”, as noted in a 1916 edition of *The Evergreen*, a weekly newspaper published by students of the Washington State College⁵.
- A second photograph depicts a class of adult students wearing headphones, concentrating on written material. The caption reads: “Learning the melody of the language”.
- Another image captures a young man standing in front of a gramophone, with Frank C. Chalfant himself standing nearby.
- The final photograph shows a small group of students wearing headphones and examining written documents, accompanied by the caption: “Analyzing the accent”.

The visual documents clearly illustrate Chalfant’s application of phonetics for didactic purposes—a pedagogical approach shaped by the training he received at the CPEE in Grenoble under Théodore Rosset. In fact, Chalfant explicitly references this educational perspective in the introduction to that issue of *The Chinook*:

⁵ Source: <https://content.libraries.wsu.edu/digital/api/collection/evergreens/id/4170/download>.

Unique in the history of modern language instruction in the United States has been the equipping of a language laboratory in this College, in which the illusive intonation and accent of the foreign languages are analyzed and systematically practiced (...). In this phonetic laboratory, the most modern and practical apparatus both from America and from Europe has been installed. The Phonograph, with sound-reservoirs and hearing-tubes, and the Dictaphone, with special attachments, serve to supplement the work of the instructor and to bring to the ears of the students the faultless accent and perfect intonation of famous elocutionists, actors and vocalists of foreign languages.

(Chalfant, *Chinook*, 1916, p. 54)⁶

Figure 5 (WSU archives) shows Frank C. Chalfant teaching in a classroom equipped with various phonetic instruments. On the wall behind him are diagrams, including a sagittal section of the phonatory organs and visual representations of articulatory movements for German sounds. In one of the photographs, a woman is seated in front of a Dictaphone, illustrating the practical, hands-on nature of the phonetic training.

One particularly revealing detail appears in another image (Figure 6), which shows that Chalfant also taught military personnel. In 1918, while serving as Head of the Department of Modern Languages at the State College of Washington, he was “assigned to the training of wireless telegraph men in the summer soldiers school for the four months beginning June 15” (WSU Archives). This suggests that Chalfant’s expertise in phonetics and auditory training had applications beyond traditional language instruction, extending into military communication and technical training during wartime.

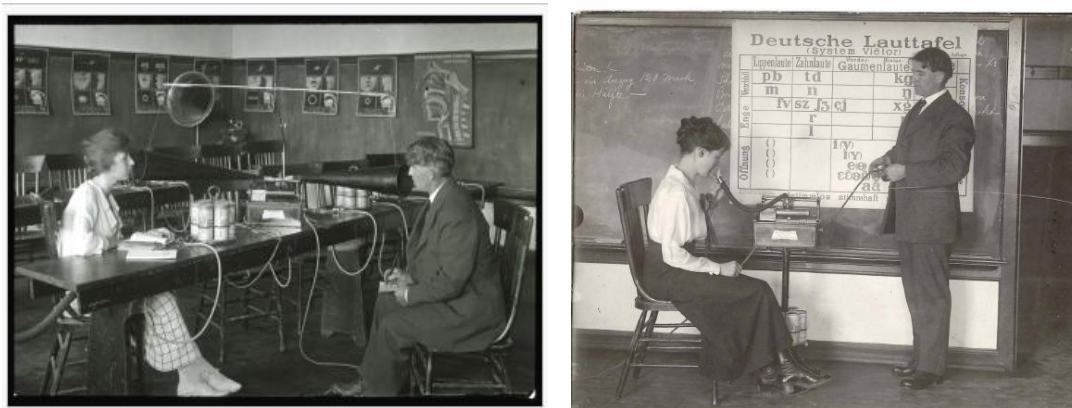


Figure 5. A. (left): Frank C. Chalfant in his language laboratory⁷ B. (right): Frank C. Chalfant with a student talking in a dictaphone⁸

Frank C. Chalfant placed significant emphasis on pronunciation in language teaching, particularly on the quality of teachers’ spoken production. This is reflected in two letters preserved in the WSU archives, dated May and June 1911, from the Seattle Public Schools Board. The letters strongly protest an inspection carried out by “High School Inspector, Professor Chalfant” at one of the city’s public schools. His evaluation of a German teacher was deemed “unsympathetic and unfair”. According to the Board, “The only criticism I secured from him was that she was weak along the lines of phonetics and pronunciation”.

⁶ Source: <http://content.libraries.wsu.edu/cdm/ref/collection/chinook/id/10821>, page 54.

⁷ Manuscripts, Archives and Special Collections, Washington State University Libraries. Historic Photographs Subject File collection, PC 4, Box 29 folder titles Laboratories – Broadcast and recording.

⁸ Manuscripts, Archives and Special Collections, Washington State University Libraries. Historic Photographs Subject File collection, PC 4, Box 29 folder titles Laboratories – Broadcast and recording.



Figure 6. Frank C. Chalfant with a student-soldier⁹

By 1916, when discussing his instructional practices, Chalfant began referring to his facility as a “language laboratory”—a term that would soon become standard. At the time, there was an ongoing scientific debate in the United States (and later in Europe), between the terms “*phonetics laboratory*” and “*language laboratory*”. These labels reflect differing educational goals: the “*phonetics laboratory*” focused on the study of phonetics as an academic discipline, while the “*language laboratory*” emphasized the didactic use of phonetics in teaching foreign languages.

In American literature on the history of language laboratories, this terminological distinction has been a subject of scholarly discussion. According to [13] (2024: 524), Chalfant was the first to use the term *language laboratory* as early as 1916, particularly within modern language teaching circles. This historical detail corrects earlier attributions that credited R.H. Waltz with introducing the term when he established language laboratories at the University of Utah in 1919, and later at Ohio State University.

2.3 Soldiers and Officers: a specialized audience at the CPEE

Alongside civilian participants, soldiers and officers also attended CPEE summer courses in the early 20th century. Although such attendees were relatively rare before World War I, a few of them went on to have distinguished military careers. The CPEE registration records mention Robert Richardson, then a young officer at West Point Military Academy, who enrolled in the summer of 1908 at the age of just 26. He would later become General Richardson with a prestigious career in the U.S. Army. In 1909, Park Powel, Master of Language at the Chicago Military Academy, attended Grenoble summer courses, and in 1913, the CPEE welcomed another notable figure: Cornelius De Witt Wilcox, Lieutenant-Colonel at the West Point Military Academy, and the author of renowned works, including a French-English lexicon. All three men studied under Théodore Rosset and benefited from his modern, experimental phonetics instruction.

At the end of World War I, Grenoble, like many French universities, hosted hundreds of American soldiers and officers. The CPEE annual report of 1919-1920 highlights the example of “Mr. WABNITZ, a student soldier at the University of Grenoble from March to June 1918, who devoted himself with admirable dedication to making our work known among his compatriots”.

The war’s end marked two significant shifts: the disappearance of German students from French universities, and the large-scale arrival of American military students. As early as 1917-

⁹ Manuscripts, Archives and Special Collections, Washington State University Libraries. Historic Photographs Subject File collection, PC 4, Box 29 folder titles Laboratories – Broadcast and recording.

1918, CPEE records reflect the emergence of "a new pedagogy designed for demobilized Americans wishing to study in France". In the words of the CPEE Annual Report dated November 20, 1918:

The demobilization of the U.S. Army, which, according to U.S. government forecasts, will take place over a fairly long period of time, will leave several thousand young people in France, for a year or eighteen months, who have either already begun university studies, or intend to devote themselves to them. It is the intention of the American government to take advantage of these young people's stay in France to have them pursue their higher education in our country, and to introduce them to French university teaching methods.

(CPEE Annual Report, November 20, 1918).

According to CPEE records, nearly 12,000 American soldiers and officers were awaiting repatriation after the war's end. Approximately 400 student-soldiers were based in Grenoble. These students were grouped by French language proficiency (elementary, intermediate or advanced) and followed an intensive program of forty hours of instruction per week. The CPEE Annual report notes that "the practical side of the teaching" was overseen by Théodore Rosset, supported by teachers from local primary and secondary schools. A formal reception party was organized in April 1919 to honor this distinguished group.

The participation of soldiers in language instruction raises important questions in the history of language teaching methodology. The system later established by the U.S. Army—known as the *American Speech Training Program* (ASTP) [12]—is often seen as a turning point. It introduced the widespread and systematic use of sound media for educational purposes during the magnetic tape era. Scholars generally agree that modern foreign language teaching in Europe, particularly in France, benefited significantly from ASTP methodology, particularly in the growth of the language laboratory [11], closely tied to the development of the audio-oral method.

However, when we consider the fact that phonetics had been used for didactic purposes in language laboratories as early as the beginning of the 20th century—by pioneers like Rosset in France and Chalfant in the U.S.—it becomes clear that the ASTP was more of a continuation than a revolutionary departure. The program likely followed a path already forged by earlier innovators on both sides of the Atlantic.

This nuanced view is presented in the *Handbook of Research for Educational Communications and Technology* [13], which outlines diverging perspectives among scholars—both defenders and critics of the ASTP's role in the development of language laboratories. As [3: 889] pointedly stated, "The Army Specialized Training Program did not, as is so widely believed, pioneer language laboratories." Scholars also highlight earlier instances of using the phonograph for language teaching in the U.S., dating back to the late 19th century. [5] documents experiments from the early 20th century that tested the efficacy of audio-based learning with American student cohorts (see also [16]).

3 Conclusion

The CPEE archives provide remarkable insight into early transatlantic knowledge exchange. In the tradition of Passy, Rousselot, and Brunot, Théodore Rosset made a significant—though often underrecognized—contribution to the practical teaching of phonetics and pronunciation for educational purposes, as well as to the development of the language laboratory. His work came at a time when few phoneticians were active in French academia [8, 9].

Other linguists, such as Pierre Delattre, further solidified this transatlantic bridge. Delattre moved to the United States in 1924 to complete his doctorate while simultaneously teaching at several American universities [10]. This mutual exchange was undoubtedly facilitated by the fertile ground already prepared for integrating phonetics into French instruction in the U.S., as

evidenced by reports of innovative practices dating back to the late 19th century. Notably, the *French Review* [16] documents the long history of French language teaching in the United States, reinforcing the idea that the developments during and after World War I were the continuation of a deeper historical trend.

4 Acknowledgments

I am grateful to Mr. Mark O'English, Archivist at Washington State University, for his invaluable assistance in researching the career and influence of Frank C. Chalfant, including access to archival materials and his generous support.

References

- [1] ABOU HAIDAR, L.: *Ce que la phonétique doit à la didactique : l'hypothèse du chaînon manquant*, *Repères DoRiF, Variations terminologiques et innovations lexicales dans le domaine de la biodiversité et du changement climatique*, n°30, <https://www.dorif.it/repere/laura-abou-haidar-ce-que-la-didactique-doit-a-la-phonetique-lhypothese-du-chainon-manquant/>, 2024.
- [2] ABRY, D., BOE, L.-J., RAKOTOFIRINGA, H.: *Théodore Rosset et l'établissement de la phonétique expérimentale à Grenoble. Une consonance entre une nouvelle discipline et le dynamisme d'un comité de patronage*, *Documents pour l'Histoire du Français Langue Etrangère ou Seconde*, vol. 20, pp. 64-72, 1997.
- [3] BARRUTIA, R.: *The past, present and future of language laboratories*, *Hispania*, vol. 50, pp. 888-899, 1967.
- [4] BOË, L.-J., VILAIN, C.-E.: *Un siècle de phonétique expérimentale*, ENS Editions, 2010.
- [5] CLARKE, C.C.: *The phonograph in Modern Language Teaching*, *Modern Language Journal*, vol. 3, pp. 116-122, 1918.
- [6] CPEE - COMITE DE PATRONAGE DES ETUDIANTS ETRANGERS, RAPPORTS ANNUELS, Archives de l'Université de Grenoble, 1896-1910.
- [7] FAVIER, R.: *Le roman de l'université. Grenoble 1339-2016*, Editions PUG & UGA Editions, 2017.
- [8] GALAZZI, E.: *Phonétique / Université / Enseignement des langues à la fin du XIX^e siècle*, *Histoire Epistémologie Langage*, vol. XVII/1, pp. 95-114, 1995.
- [9] GALAZZI, E.: *Le son à l'école*, *Centro di Linguistica dell'Università Cattolica, La Scuola, Brescia*, 2002.
- [10] GALAZZI, E.: *Pierre Delattre (1903-1967) : un phonéticien-pédagogue entre le vieux et le nouveau monde*, *Beiträge zur Geschichte der Sprachwissenschaft*, pp. 257-267, 2010.
- [11] GERMAIN, C.: *Evolution de l'enseignement des langues : 5000 ans d'histoire*, Paris, CLE International, 1993.
- [12] PALMER, R.R., WILEY, B.I., KEAST, W.R.: *The Army Specialized Training Program and the Army Ground Forces*, *Historical Section Army Ground Forces Historical Division*, pp. 28-39, 1948.
- [13] ROBY, W.B.: *Technology in the service of Foreign Language Learning: the case of the language laboratory*. In: D. H. JONASSEN (ed.), *Handbook of research on educational communications and technology*, Taylor and Francis, pp. 523-541, 2024.
- [14] ROSSET, T.: *Exercices pratiques d'articulation et de diction pour l'enseignement de la langue française aux étrangers*, Gratiery, Grenoble, 1905.
- [15] ROSSET, T.: *Recherches expérimentales pour l'inscription de la voix parlée*, Armand Colin, 1911.
- [16] WATTS, G.: *The teaching of French in the United States: a History*, *The French Review*, vol. xxxvii, 1963.

VOICELESS EVIDENCE: EXPERIMENTAL PHONETICS AND THE CONTRADICTIONS OF LINGUISTICS BETWEEN NORTHERN NAMIBIA AND COLONIAL HAMBURG

Ulrike Kiessling

University of Basel, PhD student SNF research project *The In/Audible Past*

u.kiessling@unibas.ch

Abstract: This article examines how experimental phonetics contributed to – and occasionally subverted – the colonial classification of African languages, focusing on the case of Oshiwambo in northern Namibia and its scientific construction at the Hamburg Colonial Institute at the turn of the 20th century. Based on archival research and sound recordings from 1913 and 1953/54, it analyzes laboratory measurements by Giulio Panconcelli-Calzia and field recordings by Ernst Dammann under apartheid rule.

Phonetics, often regarded as a neutral science, was deeply embedded in colonial and missionary infrastructures. The article explores how acoustic inscription, institutional assumptions, and informant agency shaped both the production and interpretation of linguistic data. While phonetic research aimed to stabilize ethnic-linguistic boundaries, it also revealed the fluidity of orders of speech and belonging – undermining rigid classificatory regimes.

By tracing how the Oshiwambo cluster emerged as a phonetic object, the study shows how recording techniques and speaker selection affected what counted as linguistic knowledge. It offers a media-historical perspective on the interplay between sound, technology, and power in colonial knowledge production.

1 Introduction

In contemporary Namibia, language and ethnic identity remain politically consequential – the link between linguistic affiliation and social belonging continues to shape public discourse. Where does this connection originate? How did language come to function as such a powerful marker of identity – and through which institutional and epistemic frameworks was this relationship constructed? These questions first emerged for me while participating in an interdisciplinary and international research project on sonic archives in southern Africa, *The In/Audible Past*. Initially focused on auditory practices and acoustic media, I encountered a set of phonetic experiments and sound recordings that pointed to a much older, deeper entanglement of language, listening, and classification and unexpectedly led back to the early history of experimental phonetics in Hamburg.

This article argues that phonetic research, while often framed as technically objective, played a formative role in shaping the classificatory logics that later underpinned colonial and apartheid policy. Yet the same research also carried within it epistemic tensions and ambiguities that could – and at times did – destabilize those very frameworks. My aim is to set a first attempt to understand how they functioned within specific historical and institutional settings, and how their legacies continue to shape our understanding of linguistic difference today. I approach this topic as a historian of media and a media theorist, aware of the fact that different forms of knowledge production are at stake – and I hope that this work can contribute to an open, critical dialogue across disciplines.

The legacy of Pierre-Jean Rousselot, who founded the first phonetic laboratory in Paris in 1889, is central to understanding these tensions. His vision of phonetics was grounded in the

physiological observation of living speech, informed by its social and cultural variation. In his early work on articulatory change in his own Occitan-speaking family, Rousselot rejected linguistic hierarchies and instead emphasized unconscious, embodied differences among speakers of the same first language.

However, this epistemic openness did not remain intact. As phonetic devices – such as kymographs or nasal glass olives in the early period, and later portable suitcase magnetic tape recorders – migrated into colonial research infrastructures, their application shifted. No longer employed solely to explore phonetic variation, these instruments increasingly served to identify, fix, and classify linguistic and ethnic difference. Over time, these tools and devices evolved – from bodily inscription to magnetic tape – reshaping not only research methods but also the epistemic distance between researcher and subject. This transformation can be traced through two key episodes: Giulio Panconcelli-Calzia's 1913 laboratory experiments on voiceless nasals at the Hamburg Kolonialinstitut, and Ernst Dammann's field recordings of Oshiwambo speakers in northern Namibia in the 1950s. Both projects were shaped by the intellectual legacy of Carl Meinhof, whose influence extended from phonetic science to the ethnolinguistic frameworks that, as scholars such as Sara Pugach have shown, would later inform apartheid policy. The article examines how techniques of inscription became entangled with the classificatory ambitions of empire – and occasionally exceeded or contradicted them. Grounded in archival and media-historical analysis, it investigates how phonetic evidence – traced across four decades – helped shape what was recognized as linguistic identity.

2 Infrastructures of Phonetic Knowledge in Colonial Contexts

In 1894, Carl Meinhof began to explore a so-called secret drumming and whistling language that would appear in a variety of regions and forms and would be spoken using a wooden instrument, or whistled or imitated with the mouth, resulting in “words spoken in a dull, murmuring tone, as a series of signals that are only known to the initiated.” [1, p. 119] Meinhof refers to Njo a Dibone, a prince, chancellor and interpreter for the governorate in Cameroon [2], whom he had met in Stettin, Germany [3, p. 49], and who drew his attention to the drum language. African academic assistants were, however, integral to the learning process at the Hamburg Institute, just as for the establishment for African studies as a discipline itself. From early on Njo Dibone, Meinhof's first African teacher, assisted with anthropological and literary pursuits. [4, p. 155]

In order to better study African languages, however, Meinhof wanted to make use of the latest recording technologies. Meinhof, who at the time was employed in Berlin as a professor of languages at the Seminar for Oriental Languages of the Royal Friedrich Wilhelm University, asked the director of the institute before his first trip to Africa in 1903 to cover the costs “for a phonograph together with 100 cylinders for 200 marks,” but his request was rejected. [5, p.81] Upon returning from the trip, he immediately sought to obtain “a room with the necessary instruments for scientific phonetics,” but this too was denied. [6, p.106/107] The search for a means of describing sound beyond semantics, as well as the demands of Meinhof's contemporaries for methodological rigor, ultimately led to the foundation of the Phonetic Laboratory following his appointment in Hamburg. The Phonetic Laboratory in Hamburg thus emerged from a collaboration between the Hamburg Colonial Institute and Meinhof's experimental phonetics. The laboratory was headed by Giulio Panconcelli-Calzia. By 1933, there were approximately 60 chairs, lectureships, and institutes for experimental phonetics in Europe. [7, p.114] Meinhof's push for recording technologies was also aligned with broader epistemic shifts in phonetics more generally. Across Europe, researchers sought to move beyond impressionistic transcription toward reproducible, instrument-based observation. One of the most influential figures in this movement was Pierre-Jean Rousselot, whose work laid the theoretical and technical foundations for experimental phonetics as a modern science.

In the preface to his *Principes de phonétique expérimentale* (1897) [8, pp.14-16], Pierre-Jean Rousselot distinguishes between two traditions of phonetic inquiry: historical phonetics and descriptive phonetics. Historical phonetics, he argues, emerged first – driven by the reconstruction of sound change through comparative study and textual analysis. Descriptive phonetics, by contrast, focuses on the observation of living speech. For Rousselot, the two are deeply interdependent: one cannot interpret historical sound shifts without a physiological understanding of how speech is actually produced. Yet it is descriptive phonetics, he insists, that opens the door to experimentation. Dissatisfied with the limits of the unaided ear, Rousselot turns to instruments and inscription devices to measure airflow, articulatory motion, and acoustic variation. This methodological shift – from listening to recording, from impression to inscription – marks the transformation of phonetics into a modern experimental science. Its goal is not merely to register what is heard, but to construct reliable knowledge formations and robust categories through controlled observation and reproducible data. In this framework, language becomes a physiological and measurable phenomenon – accessible through instrumentation.

This redefinition of phonetics as a materially grounded science shaped the research of the Hamburg Phonetics Laboratory, including the early work of Giulio Panconcelli-Calzia. In his 1913 experiment on voiceless nasals, Panconcelli-Calzia employed tools such as kymographs, glass olives, and stylus-based inscription devices – not in the service of linguistic typology, but as part of an open-ended investigation into the articulatory properties of speech. His approach still resonates with Rousselot’s vision: a form of experimental curiosity oriented toward discovery rather than verification.

Four decades later, in the work of Ernst Dammann, the role of instrumentation had changed. While technical methods such as kymography remained in use and the technology of and tape recording was added to the arsenal of apparatus, their epistemic function had become more standardized and utilitarian. Rather than exploring the boundaries of articulatory possibility, Dammann’s recordings aimed to document, segment, and classify speech in ways that aligned with administrative needs – particularly in the context of apartheid-era linguistic mapping. The experimental dimension, once central, receded into background routine; method became tool, not question. What began with Rousselot as a methodological project rooted in uncertainty and embodied observation had become, in Dammann’s hands, an instrument of institutional classification. This transformation – subtle but consequential – marks a shift not only in scientific practice, but in how phonetic knowledge was imagined, operationalized, and politically mobilized.

The scholarly environment of the Phonetic Laboratory raises questions of the entanglement of colonization, phonetics and sound studies, including questions concerning the epistemological conditions of recording techniques. This opens up a set of systematic and infrastructural questions: Which institutions and laboratories and which recording techniques were involved? How does this relate to the classification and categorization of languages as colonial practice? What methods produced an archive predicated on linguistic boundary-drawing? How do these different forms of knowledge intersect within a Namibian colonial context? The Hamburg Scientific Foundations drew their foundational capital significantly from colonial activities in Africa; Meinhof’s work there fell directly under the supervision of the Prussian Minister of Finance. [9, p. 7, 74] In 1910, the Institute for Colonial Languages in Hamburg and the Laboratory began to collaborate. From this collaboration, a network emerged dedicated to phonographic recordings and the simultaneous study of their technical, material and sonic conditions.

The roots of experimental phonetics at the Kolonialinstitut Hamburg lie at the intersection of colonial governance, missionary linguistics, and laboratory science. Carl Meinhof, being involved in the project of Bantu studies, tried to explain not just the origin of language but the origin of being and culture at large. Therefor he played a key role in institutionalizing African linguistics within the structures of the German empire. As a professor at the Colonial Institute

in Hamburg, where Dammann was a student, Meinhof supervised both missionary and academic linguistic research. Giulio Panconcelli-Calzia, a trained phonetician, introduced Rousselot's approach of bodily techniques and mechanical inscription devices, such as the kymograph, to register and analyze speech.

This shift from experimental openness to procedural standardization reflects not a decline in scientific rigor, but a transformation in the role of method itself. In Panconcelli-Calzia's time, phonetic instrumentation was still part of an evolving epistemic field: methods were actively tested, discussed, and debated. By the time of Dammann's work, however, experimental techniques had become routine – technically refined, yet less often subject to methodological reflection.

3 Phonetic Bodies, Colonial Frames: The 1913 Nasality Experiment in Hamburg

The institutional roots of Panconcelli-Calzia's 1913 nasality study lie in late nineteenth-century France, particularly in the work of Pierre-Jean Rousselot. Barred from doctoral study in Germany due to his partly Italian nationality, he moved to Paris to study with Rousselot at the Institut de laryngologie et orthophonie. There, in 1904, he completed a dissertation titled *De la nasalité en italien*, marking the start of his lifelong engagement with nasality as both a physiological and phonetic phenomenon. Rather than aligning phonetic analysis with emerging national language norms, Rousselot used experimental methods to study how speakers diverged from standardized forms. In doing so, he implicitly challenged the assumption that linguistic legitimacy resided in fixed, written standards. In his dissertation *Les modifications phonétiques du langage étudiées dans le patois d'une famille de Cellefrouin*, Rousselot analyzed articulatory change within his own Occitan-speaking family across five generations. He investigated unconscious variation among first-language speakers using kymographs, nasal pressure tubes, and laryngeal sensors. His goal was epistemic: to render perceptible what the human ear alone cannot reliably detect. As Didier Demolin shows, Rousselot considered the phonetic instruments he developed – such as kymographs and palatographic devices – not merely as tools, but as epistemic extensions of perception, capable of registering unconscious or imperceptible articulatory variation that escaped the human ear. [10]

The early years of the Hamburg phonetic laboratory were characterized by a hybrid orientation: on the one hand, a commitment to technical experimentation and on the other, increasing entanglement with the classificatory ambitions of German colonial linguistics under Carl Meinhof. Panconcelli-Calzia, after returning to Germany, conducted experimental phonetic studies on African languages at the Hamburg Phonetic Laboratory. In 1910 he was appointed by Carl Meinhof as a phonetics assistant at the newly founded Phonetic Laboratory of the *Hamburg Colonial Institute*, where he became associated with leading phoneticians and speech and voice pathology researchers, as Hermann Gutzmann. There, he combined acoustic technology with colonial language research, and the issue of nasal sounds a focus of his experimental investigations. His interest in nasality thus emerged from the intersection of physiological phonetics, technological experimentation, and the classificatory ambitions of colonial linguistics. Panconcelli-Calzia, trained in the Rousselotian tradition, brought with him an interest in the physiological dimensions of nasality rather than in overt ethnic or racial classification. His experiments on voiceless nasals – applied later to Oshindonga – reflected this initial focus. However, the colonial research infrastructure into which his work was embedded increasingly steered phonetic data toward taxonomic ends. While Panconcelli-Calzia built on Rousselot's physiological methods, his experiments increasingly took on a classificatory function within the colonial research infrastructures of German African studies. What began as a descriptive science of variation thus gradually became a tool in the construction of linguistic and social difference under imperial rule. As Sara Pugach has shown, Carl Meinhof's linguistic theories left a lasting

imprint on the work of Nicholas van Warmelo [11], whose ethnolinguistic classifications later informed aspects of apartheid policy in South Africa. Meinhof's influence, however, extended beyond South African administration: "The tale of Meinhof and his German colleagues, who shaped and founded the discipline of African language and culture studies – Africanistic – is nonetheless critical for understanding the history of racial thought in Germany and how that thought contributed to segregationist and apartheid ideology in South Africa." [12, p. 2]

On 14th of April 1913 the son of Martti Rautanen, a Finnish missionary stationed in northern Namibia, happened to be in Hamburg under unknown circumstances yet to be examined. He was asked to participate in an experiment focused on the production of voiceless nasality, a phonetic feature common in Oshindonga. Using a kymograph, Panconcelli-Calzia recorded airflow and articulatory movement directly on the subject's body – placing tubes in the nostrils and mouth to capture nasal airflow: "I guided the sound waves coming out of the mouth through a mouth funnel into a writing capsule, which touched the sooty paper of a rotating drum; at the same time I examined the processes in the nose through a glass olive inserted into a nostril, which was also connected to a writing capsule." [13, p. 257] These procedures illustrate the intense corporeality of early phonetic science. In terms of my own disciplinary perspective as a media-historian, it seems striking that Panconcelli-Calzia did not show interest in the sociolinguistic context of the language he examined: «[Carl Meinhof] took Mr Rautanen (sic!) from the Finnish Missionary Society to the Phonetic Laboratory to let me hear some peculiar Ndonga sounds and possibly have them examined objectively. Mr Rautanen was, as far as I know, born and educated in South Africa, so I consider him to be a reliable source in phonetic terms.» [14]

In his publication, Panconcelli-Calzia vaguely referred to the language as "originating from South-West Africa," without specifying its regional or social affiliations. His test subject, Reinhold Rautanen - 36 years old at the time of examination - was identified by Calzia as the missionary Rautanen who had grown up in "South Africa", thus confusing father and son, South Africa and South West Africa. These imprecisions reveal epistemic gaps in the otherwise meticulous, precise and accurate work of Panconcelli-Calzia. They point to a lack of communication between the test subject and the interrogator and the production of colonial effect: epistemic blind spots embedded in colonial classificatory routines of scholarly research. As Panconcelli-Calzia reflects in his later writings, the experimental phonetician abstracts from all language: "For him, there are only beings equipped with respiration, larynx, and vocal tract who phonate." [15, p. 252] This epistemic logic – methodologically rigorous, yet detached from any semantic or sociological context – made it possible to treat voices not as acts of communication but as acoustic signals. In his analysis, Panconcelli-Calzia referred to various missionary linguists before him concerned with these matters, such as Brincker, who had dealt predominantly with Oshikwanyama, another Oshiwambo dialect. Soon, Panconcelli-Calzia found that he had to contend with a fundamentally new quality of sounds that required new experimental conceptualizations. He called these "snap sounds", and tried to grapple with these phonetic qualities in Oshiwambo languages at a distance by referring to Hamburg's collections on isiZulu and Korrana and their click sounds. He then attempted, following Brincker's lead in describing nasal sounds for Oshikwanyama, to exclude the sounds of clicks and snap sounds as observed in Rautanen's speech samples:

«If the sounds described by Brincker are the same as those I mean, it is certain that they have nothing in common with clicks. Indeed, a writing apparatus records snap sounds, which are known to be produced by sucking movements in the mouth, through a downward deflection [...]; if the sounds in question were nasal snaps, the negative deflection would be found on the nose line.» [16, p. 260]

Panconcelli-Calzia concluded through experimentation in Hamburg that the "examined informant for Ndonga" exhibited both voiced and voiceless nasal sounds. This underscores his primary interest as a researcher of nasality – namely, demonstrating the existence of voiceless

nasal sounds. Presumably, his collaboration with Carl Meinhof, who assumed that voiceless nasal sounds evolved from voiced ones, informed this endeavor. However, it is pertinent to note here that in early experimental phonetics, ethno-linguistic classifications were still under negotiation, and there were no finalized conceptualizations. The case of the Ndonga study involving the Finnish missionary's son Rautanen as a test-person highlights the linguistic and phonological interest of scholars in sound phenomena rather than pre-conceived ideological conclusions or ethnic projects. Damman implicitly applies the paradigms developed by Panconcelli-Calzia in his experimental design without referring to its genealogies. Damman's later work then can be read as a return to Panconcelli-Calzia's concern with voiceless nasals – but transposed from Occitanian to Namibian languages, from the laboratory to the colonial field, from sooty paper to magnetic tape.

4 Fieldwork, Tape, and Feedback – From Phonetic Traces to Apartheid Taxonomies, Dammann's Recordings (1953–54) in Context

Four decades after Panconcelli-Calzia's laboratory-based experiments, Ernst Dammann returned to the question of voiceless nasals – this time through field recordings in apartheid-era Namibia. While still shaped by Meinhof's classificatory vision, Dammann's work reflected both a shift in medium and a new kind of field-based engagement. Technologically, the shift from kymograph to magnetophon marked a move from body-based inscription to mediated acoustic fidelity and portability. The spatial relation between researcher and subject changed as well; microphones allowed for more distanced recordings, reflecting broader transformations in fieldwork practices. In 1953, Dammann traveled to Namibia equipped with a magnetophone to record spoken examples of what he referred to as Ndonga, and to investigate the phenomenon of voiceless nasality – once again inspired by Meinhof's earlier hypotheses, that such nasal sounds represented a secondary development from voiced nasals in the historical phonology of Bantu languages. It remains unclear whether Panconcelli-Calzia and Dammann ever discussed this topic. Dammann's expedition was funded by the DFG by the claim that this region remained an underexplored linguistic area – demonstrating the continuity of a German academic desire to “discover” linguistic Others and probably to reclaim authority over a former colonial territory. Dammann explicitly referenced Panconcelli-Calzia's earlier work and sought to build on it. [17] Building on a disciplinary lineage that sought to render speech scientifically observable, Dammann continued the attempt to replace auditory impression with measurable data. Yet whereas earlier phoneticists, such as Panconcelli-Calzia, “armed the ear with instruments” to record the body's vibrations [18, p. 248], Dammann relied on magnetic tape – a shift from immediate bodily inscription to mediated acoustic analysis.

Prior to embarking on their 11-week journey into Ovamboland, the Dammanns were equipped by the Phonetic Laboratory Hamburg with a Grundig TK 40 [19] magnetophone and BASF tapes. The specific instructions regarding the operation of the magnetophone are not documented, but it is most likely that Ruth, wife and unpaid assistant of Ernst Dammann who was responsible for technical matters, received instructions on its operations from Mr. Cremer, the technician from the Phonetic Laboratory. The portable magnetophone, although cumbersome in a practical travel case, was more easily transportable, compared to earlier recording devices. Ruth, who mostly carried the magnetophone herself to protect it, gave detailed descriptions of this transportation process in her diary entries: "I have the magnetophone on my lap, it's a miserable ride on the paths and I have to be very careful with it. If the magnetophone is not hit by the bumps on the road, it is because my arms are." [20] The tape speed ensured high-quality sound and speech recording, efficient tape consumption, and optimal playback quality: "Special design features ensured the highest tracking accuracy and greatest operational reliability," with a total playing time of up to 16 hours. [21] However, the weather conditions in Oniipa

affected the recording quality, as Ruth described: "The diesel motor is started up, but unfortunately it doesn't give a regular current, so the playback of our tapes is uneven and not at the right pitch. Maybe it will get better if I switch the machine to 110 V, but I can't do that today. In any case, it's not broken." [22]

While Ernst Dammann was a research assistant at the "Seminar for African and South Sea Languages at the University of Hamburg" from 1930-33, Carl Meinhof had made him aware of manuscripts in the Ndonga language, a Bantu language in South West Africa. [23] These manuscripts were part of the personal estate Martti Rautanen. The editing and orthographic adaptation of the Rautanen manuscripts in the 1950s had not been made by missionary scholars, however, but by local speakers, translators, and intellectuals. These had played a decisive role in shaping what would become standardized Oshindonga. In their 1975 anthology, Dammann and Tirronen, a Finnish teacher and missionary, explicitly acknowledged that the publication would not have been possible without the assistance of local "guarantors" – that is, native speakers who provided validation, interpretation, and linguistic judgment. The linguistic insight of informants such as Timothy Namuhuya, who translated, mediated, and helped record the oral narratives of his brother H.D. Namuhuya, widely known as the first Oshiwambo speaking published poet. Other informants included Hosea Nampala and Jairus Mbenzi, who structured many of the collected materials. Their work exemplifies the entanglement of local expertise and colonial knowledge systems. The standardization of Oshindonga reveals a layered field of linguistic negotiation, shaped by asymmetries of power but also by shared epistemic labor. At the outset of the Phonetic Laboratory's work in Hamburg, the center functioned primarily as an experimental space. Over the years, standardized and validated processes were introduced, which the Dammanns relied on when they embarked on their journey to Namibia in 1953.

In his 1957 article "*Stimmlose Nasale im Ndonga*" [24] published in the journal *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*, Ernst Dammann refers to the experimental study conducted by Panconcelli Calzia in 1913 [25] at the Phonetic Laboratory. Calzia, who never travelled to Namibia himself, acquired his knowledge from circulating journals and verified the information through experimentation in Hamburg. To understand how Dammann positioned himself within this tradition, it is worth revisiting Panconcelli-Calzia's original methodology.

Panconcelli-Calzia's phonetic apparatus connects a set of human and non-human elements. The airflow emanating from the lungs of a person, interrupted rapidly by the vocal cords, produced a measurable sonic structure, or sound. This sound, in turn, vibrated through the writing apparatus used for examination, causing the membrane connected to a writing lever to leave small peaks impressions on the paper. Thus, the visual and graphical peaks made visible through the kymogram obtained from the experiment with a native speaker — or pseudo-native speaker such as Rautanen — indicated the vibrations and the voicing of the sounds by representing them as inscriptions on paper. In this procedure, the experimental participant, the person, tends to vanish from the scene and the acoustic fact of "voicing" remains. The absence of diagrammatic peaks in the inscription then suggests the absence of voicing, indicating voicelessness. Rautanen, the missionary informant, uttered words such as "*ombu – onde – ondja – ondya – ompo – ontsi*" into the apparatus to observe the different nasal sounds. As only the acoustic evidence remained, the experiment could count as an attempt to inquire into the sounds' origins within various Bantu groups.

In this context as in others, phonetics researchers disregarded the intra-familial background that Rousselot himself had put into focus. As often, the social and linguistic context of informants was left out in order to fit experimental recordings into a classificatory system or in order to explore sonic roots and origins of language. This was a usual procedure in phonetic recordings. Yet in 1950s Namibia, multilingualism, natal communities, and social positioning were all central to negotiations of belonging and identity, particularly in the lead-up to Apartheid.

Speaking a particular language could mark a person's inclusion in – or exclusion from – specific social groupings, depending on the political context. While those criteria were naturalized in the process of research, they were in fact politically charged. In Dammann's case, they were deliberately ignored in the construction of classifications. These crucial aspects of the experimental person's as well as the informants' socio-formation disrupted the scientific and archival assumptions and claims of phoneticians. They practiced their research in times of hegemonical acceptance of an image produced by white supremacy, an ethno-linguistically segregated image of the social world. These tensions become audible in the recordings themselves: sociolectal features surface in Namuhuya's speech, and his social mode of speaking appears to shift under Dammann's recording conditions – suggesting a performative adjustment to the asymmetrical setting of linguistic fieldwork. This observation, which was also remarked upon by some of Namuhuya's contemporaries, points to the complex social negotiations embedded in phonetic inscription. While a fuller analysis exceeds the scope of this article, it will be further developed in my PhD project. Nonetheless, the informants left traces of their social complexity embedded in their voice samples.

We can conclude from these experiments that they were made possible only through the collaboration of local informants, whose linguistic expertise, narrative knowledge, and translational labor shaped the recordings. Figures like Timothy Namuhuya, Jairus Mbenzi, and Hosea Nampala were not simply subjects, but active co-producers of the data. Their roles ranged from storytellers and translators to linguistic validators, and their voices remained embedded in the archive – even when the text rendered them invisible.

The measurements were not made directly on a person this time but objectivity was supposed to be given even more than before via kymographs through recordings made with a magnetophone: Thus, Dammann described examining magnetophone recordings made with his wife during their 1953 trip: "On December 22, 1953, at the Finnish Mission Station in Oniipa. The informant was Timothy Namuhuya." [26] The measurements were made again later on using a kymograph, not on the direct vibrations of the body expressions of phonemes, but of on tape recorded sounds. In later reflections on this Oshiwambo dialect research, Dammann mentions only in passing alongside Namuhuya some other key individuals involved in the experimental setup back in Hamburg: the head of the Phonetic Laboratory, Otto von Essen, who interpreted the measurements made by the technician Cremers. In his research publications on Ndonga, Dammann emphasized that there had been no robust scientific studies of Ndonga in the past, marking a place for himself in his scientific milieu. For his onward studies of many Namibian languages in the phonetic laboratory in Hamburg, he had recordings made by his informants in Namibia. These were local language experts, poets, intellectuals and translators, whom he encountered during his journey, such as Hosea Nampala, Jairus Mbenzi, Timothy Namuhuya, Filemon Silvanus, Lena Fenter, to name just a few. On the one hand, he had people tell stories or record vocabulary, repeating them three times to capture different sound qualities as the meaning changes due to height of tone or length of vowels.

Both Panconcelli-Calzia's and Dammann's work contributed – albeit unintentionally—to the formation of the Oshiwambo cluster as a recognized linguistic object. Originally treated as separate languages (e.g., Oshindonga, Oshikwanyama), these varieties were gradually understood as mutually intelligible dialects within a larger cluster. This conceptual shift was a result not only of field observations but of the iterative epistemological feedback between local informants, missionary grammars, and phonetic data.

Paradoxically, while colonial and apartheid regimes sought to stabilize ethnic identities through linguistic classification, the actual data from phonetic and linguistic research often destabilized such rigid categories. The dynamic variation and dialect continua documented by researchers undermined the claim that ethnic or linguistic boundaries were fixed and could be mapped onto territorial divisions.

Under South African rule, especially during the apartheid era, linguistic classification was instrumentalized for territorial segregation. Ethnic groups were assigned specific homelands based on perceived linguistic affiliation. However, the granular detail of linguistic studies - such as Dammann's recordings - complicated this schema. The observation that speech varieties interpenetrated and that dialect boundaries were porous posed a challenge to apartheid logics.

Indeed, while Dammann's framing was aligned with classificatory ideologies, the linguistic variation captured in his recordings often contradicted them. In this sense, even as phonetic science served the logic of apartheid, it also contained the beginning of its own undoing.

5 Conclusion: Phonetics, Power, and Epistemic Ambivalence

The comparative analysis of Panconcelli-Calzia's and Dammann's research on voiceless nasality reveals how experimental phonetics both shaped and was shaped by colonial infrastructures. Rather than treating their work as purely technical, this article situates their inscription practices – kymographic tracing, nasal airflow measurement, and magnetic tape recording – within a media history of sound that foregrounds the epistemic conditions of phonetic knowledge production.

Panconcelli-Calzia's 1913 study exemplifies a moment when experimental phonetics was still epistemically open. Influenced by Rousselot, he used tools like the kymograph and glass olive to render voiceless nasals visible, translating bodily movement into visual data. His work focused on articulatory physiology, not ethnic classification, but nonetheless relied on colonial infrastructures and missionary networks. It reflects a moment when the very criteria of what should be measured were still in formation.

By the 1950s, Dammann's framing of Panconcelli-Calzia's experimental set-ups was aligned with classificatory ideologies or ideological classifications. However, the linguistic variation captured in his recordings often contradicted these. Yet it was precisely the act of rendering such variation visible and administratively legible that founded the orders of linguistic differences, on which later systems of segregation were based. While on the one hand phonetic science served the logic of apartheid, it also contained the potentials of its own undoing.

This study contributes not another general account of linguistics and colonialism, but a closer look at how phonetic inscription technologies mediated this relationship – technically and epistemically. Focusing on voiceless nasals brings into view how experimental detail became entangled in classificatory regimes. These recordings did not merely reflect linguistic structure; they helped construct the differences to distinguish what could be heard, known, and governed.

Ultimately, this comparison shows how experimental phonetics operated at the intersection of scientific ambition, colonial constraint, and linguistic complexity. While it presented itself as neutral and empirical, it was embedded in assumptions about language, race, and territory. Yet the very recordings – nasality, dialect variation, multilingualism – often undermined those assumptions. In this sense, the recordings became forms of *voiceless evidence*: shaped by asymmetrical systems, but still carrying epistemic traces that exceeded their intended purpose.

The shift from bodily inscription to acoustic capture marks a transformation in both method and the politics of listening. Informants were not passive subjects, but complex agents whose voices – though often silenced in text – remained audible in sound. Revisiting these episodes invites a more nuanced view of science in colonial contexts: not as a unified instrument of power, but as a field of contradictions, negotiations, and epistemic leakage.

6 References

- [1] MEINHOF, C.: Die Geheimsprachen Afrikas. In: *Globus*, vol. 8, p. 119, 1894.
- [2] *Ibid.*

- [3] GERHARDT, L.: Carl Meinhof – Das Leben des ersten Ordinarius für Afrikanistik, Hamburg 2022.
- [4] PUGACH, S.: Of conjunctions, comportment, and clothing: The place of African teaching assistants in Berlin and Hamburg, 1889–1919. In: *Ordering Africa: Anthropology, European Imperialism and the Politics of Knowledge*, Manchester 2011.
- [5] GERHARDT, L.: Carl Meinhof – Das Leben des ersten Ordinarius für Afrikanistik, p., 7, 74, Hamburg 2022.
- [6] *Ibd.*
- [7] MEHNERT, D., M. PÉTURSSON, R. HOFFMANN: *Experimentalphonetik in Europa*. Institute, Wissenschaftler, Leistungen, Dresden 2016.
- [8] ROUSSELOT, P.-J.: *Principes de phonétique expérimentale*, vol. 1, Paris, 1897.
- [9] GERHARDT, L.: Carl Meinhof – Das Leben des ersten Ordinarius für Afrikanistik, Hamburg 2022.
- [10] DEMOLIN, D.: *Rousselot et l'étude des langues du monde*. Manuscript, Laboratoire de Phonétique et Phonologie, Université Sorbonne-Nouvelle, n.d. (unpublished typescript), section 5, Discussion.
- [11] PUGACH, S.: Carl Meinhof and the German Influence on Nicholas van Warmelo's Ethnological and Linguistic Writing, 1927-1935. In: *Journal of Southern African Studies*, vol. 30, No. 4, Special Issue: Writing in Transition in South Africa: Fiction, History, Biography, pp. 825-845, 2004.
- [12] PUGACH, S.: *Africa in Translation, A History of Colonial Linguistics in Germany and Beyond, 1814-1945*, University of Michigan Press, 2012.
- [13] PANCONCELLI-CALZIA, G.: Objektive Untersuchungen über die stimmlosen Nasale im Ndonga. In: *Zeitschrift für Kolonialsprachen*, vol. 6, 1915/1916.
- [14] *Ibd.*
- [15] PANCONCELLI-CALZIA, G.: Das Als Ob in der experimentellen Phonetik, *Annalen der Philosophie und philosophischen Kritik*, vol. 4, no. 4/5, 1924/1925.
- [16] PANCONCELLI-CALZIA, G.: Objektive Untersuchungen über die stimmlosen Nasale im Ndonga. In: *Zeitschrift für Kolonialsprachen*, vol. 6, 1915/1916.
- [17] DFG Archiv, Schriftgut DA 12/1-7.
- [18] PANCONCELLI-CALZIA, G.: Das Als Ob in der experimentellen Phonetik. In: *Annalen der Philosophie und philosophischen Kritik*, vol. 4–5, 1924–1925.
- [19] Dammann himself would later claim that he was using a TK Grundig 41, which was produced after his travels. That's why I am suggesting he was using the latest model at his time.
- [20] BAB PA_39_ Tagebuch Ruth, p. 31. [translation by author]
- [21] Advertisement journal Grundig_TK-40-41-42-Prospekt-1965.jpg
- [22] BAB PA_39_ Tagebuch Ruth, p. 31. [translation by author]
- [23] DAMMANN, E., T.E. TIRRONEN: *Ndonga Anthologie*, Reimers Verlag, 1975.
- [24] DAMMANN, E.: Stimmlose Nasale im Ndonga, In: *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*; later lead by G. F. Meier as ZPSK.
- [25] PANCONCELLI-CALZIA, G.: Objektive Untersuchungen über die stimmlosen Nasale im Ndonga. In: *Zeitschrift für Kolonialsprachen*, vol. 6, pp. 257–263, 1915/1916.
- [26] DAMMANN, E.: Stimmlose Nasale im Ndonga, In: *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*; later lead by G. F. Meier as ZPSK.

TEACHING RHETORICS IN HUNGARY IN THE 13TH–19TH CENTURIES

Péter Libárdi

*National University of Public Service, Budapest, Hungary
libardi.peter@uni-nke.hu*

Abstract: This study provides an overview of the nearly six hundred years of rhetorical education in Hungary, highlighting the periods and authors that had the most remarkable influence on its development. Following the principle of synchronicity, it utilizes the rhetorical knowledge of antiquity, and clarifies the chronological arc of continuity with a diachronic analysis. More deeply, it discusses the curriculum, textbooks and teaching methods of Hungarian rhetoric education.

1 Introduction

Speech is the human activity of remembering, thinking and reflecting, through human language, the most natural form of human communication. It is a skill that can be acquired as a result of years of learning and is linked to an individual's development and socialisation [20]. As Pinker [26] defines it, „Language works by having in each person's brain a dictionary and a glossary of words (mental lexicon) and a set of rules that determine how the elements of the dictionary are related and how the relationships between concepts are expressed (mental grammar)”. It is during lexical access (searching for and finding a given linguistic, speech sign) that the currently activated vocabulary is loaded [16]. Structurality becomes a function of the complexity of the grammatical figures and structures used in speech, with the production system being organised from the early years onwards according to the adult pattern. This activity is the task of the school education field, when the individual's linguistic and speech patterns are no longer acquired through environmental experience alone, but the object of reflection is the linguistic system itself, its structure and its rules of operation. The development of linguistic awareness begins in preschool and can last for the rest of our lives [18], and its increasing tendency is most strongly observed through the language use patterns of youth [21]. Its deepening and inculcation in the process of mother tongue education relies, among other things, on the framework of rhetoric teaching, based on the structural units of the speech act, its genres and types of arguments, applying the rhetorical theory and practice of classical rhetoric. Its grammar and terminology were developed by the Greek and Roman rhetoricians, creating a model of grammatical description, and passed down to later ages. This is the reason why this rhetorical vocabulary survived into the Middle Ages, but also continues to live on in today's national languages, because the rhetoric of all European languages began with the translation and adaptation of ancient rhetoric, and its terms have been mostly preserved for the sake of clarity [3, 12].

Speech is still central to the teaching of rhetoric today, as the essence of rhetoric is to learn how to communicate effectively and persuasively. Rhetoric teaches not only how to construct arguments logically and influence an audience, but also how to communicate our ideas clearly, expressively and effectively. Through the practice of speaking, students learn how to structure their message clearly and persuasively, how to use rhetorical devices (e.g. questions, repetition, similes), how to use their voice, body language and facial expressions to enhance their oral impact, and how to interact with their audience.

The introduction of rhetorical education in national languages took place in several stages, depending on historical and cultural contexts, mainly from the end of the Middle Ages through the Renaissance and the Reformation to the 18th and 19th centuries.

Bible translations (such as Luther's German translation) strengthened the status of national languages, and with this, rhetorical training gradually appeared in these languages. Thanks to educational reforms, the study of rhetoric in national languages began in schools, initially in Protestant grammar schools. Works on rhetoric were published in national languages, and works based on these were taught in German, French and later Hungarian, often serving as aids for religious speeches (sermons).

As national literatures emerged, rhetoric as the art of eloquence became increasingly linked to writing and public speaking in the native language. Handbooks on public speaking in the native language were published in education, and in Hungary, János Apáczai Csere (a Transylvanian Hungarian educator, philosopher, Calvinist theologian, and pioneer of Hungarian education and the dissemination of scientific knowledge in the native language) and other Protestant scholars played an important role in the early stages of rhetorical education in Hungarian.

The formation of modern national rhetorical traditions began in the 19th century, and by this time, rhetorical education in national languages was already widespread. As part of national identity and civic education, rhetoric became part of school education, no longer only in Latin or German, but also in the native language of the country.

Today, rhetoric has changed a lot in both space and time, but its classical foundations are still valid. It has evolved into a discipline of systematic thinking and self-expression, as well as civilized and fair argumentation and debate. As the discipline of good speech, it is independent and multidisciplinary: it has its own methodology and objectives, which also determine the methodology for discussing its concepts. It also discusses principles that appear in everyday communication processes in a broader sense. It can therefore be understood as a related discipline of text analysis, which shapes the creation of text (speech) according to the criteria of successful communication. Traditional rhetorical types and tools can be applied in the fields of modern communication theory, text analysis, and social psychology. Classical rhetorical theory and practice are revived in modern linguistic, psychological and psycholinguistic theories within the framework of text analysis, and through the realization of text, they become a form of expression in public life.

In line with the social changes of our age, rhetoric is approached in a new way as social intelligence: sensitivity and behavior, which provides individuals with the necessary skills to become members of the community, assert themselves and understand others, without in any way rejecting the knowledge and use of the classical discipline [2, 9].

2 Hungarian rhetoric and its teaching in the context of international trends

The preservation, development, and influence of rhetorical knowledge across ages is also due to its continuous presence in education. Aristotle (384–322 BC) [5] gave birth to the first scientific rhetoric with his *Rhetoric*. The first systematic rhetorical manual in Latin was Cornificius's *Rhetoric dedicated to C. Herennius* (80s BC), in which the author summarises the rhetorical system that had developed up to that point, drawing on Greek sources. Quintilian (35–100 AD) examines rhetoric in his ancient rhetorical manual *Institutio oratoria*, placing it within the context of education. From the perspective of the art of rhetoric, the latter is considered to be a timeless classic. Ancient rhetoric, especially the works of the authors mentioned above, had a decisive influence on the formation and development of Hungarian rhetoric and speech teaching [4].

In addition to rhetoric, these theoretical foundations also provided Hungarian-language education with a toolkit for argumentation, style, and persuasion. Since rhetoric was taught in Latin for a long time in Hungary, ancient authors were directly present in the curriculum: Protestant colleges, for example in Debrecen, Sárosspatak (Hungary) and Kolozsvár (now in Romania), often held so-called declamations (public speaking exercises), during which students wrote and delivered argumentative speeches and participated in debates where they had to defend opposing points of view. These speeches were structured according to the rules of ancient rhetoric: exordium (introduction), narratio (historical part, presentation of facts), argumentatio (arguments, refutations), peroratio (conclusion, impact).

The influence of medieval works on Hungarian rhetorical education was fundamental and mediating in nature: these works formed the link between the ancient rhetorical tradition and later Hungarian-language rhetoric. Although rhetoric flourished mainly in Latin in the Middle Ages, Hungarian education and rhetorical culture also benefited greatly from this, mainly through the church and schools. Rhetoric was included in the school curriculum as one of the subjects of the septem artes liberales, together with logic and grammar, as one of the trivium (grammar, dialectic, rhetoric), and was its third pillar. This system was adopted in Hungary by monastery and cathedral schools, and later by town schools. The curriculum was based on excerpts from ancient authors, medieval handbooks, and parables, and its aim was to develop rhetorical skills for preaching, debating, and correspondence.

This is evidenced by contemporary textbooks, the dictamen, which developed into three types: poetry, letter writing and preaching (the latter being a typical medieval ecclesiastical genre with no ancient precedent). Even in monastery schools, the main aim of this discipline in higher education was still considered to be administrative work and the drafting of official letters, documents, and laws.

Renaissance took over the role of Medieval rhetoric, but in a renewed and deepened form, and it became an integrative, aesthetically oriented discipline. During the Reformation and Counter-Reformation, spoken word regained its prominent role in church sermons.

From the 17th century, the relationship between rhetoric and belles-lettres, and aesthetic questions received much attention also in education. By the late 1800s, departments of rhetoric were established at several universities in the United States (for example, at Cornell University: Department of Oratory and Rhetoric in 1872; University of Chicago in 1892; Northwestern University: School of Oratory in 1895) [7]. It is important further to mention Hugh Blair's *Lectures on rhetoric and belles lettres* [8], in which the author presents an unified and logical knowledge of rhetoric, literature, and literary criticism on a philosophical and pedagogical basis. His work reached Hungary at the end of the 18th century, and was used in translations and excerpts in Latin and German, and later in Hungarian in secondary school rhetorical education, increasingly incorporating Blair's ideas on style and taste.

George Campbell's work, *The philosophy of rhetoric* [10], had an important but more indirect influence on Hungarian rhetorical education, especially in the first half of the 19th century. He discussed association, sympathy, moral considerations, and persuasion as the desired outcome of the cooperation between teaching, imagination, emotion, and will.

3 The history of rhetorical education in Hungary

One of the earliest Latin sources of rhetorical education in Hungary is the *Esztergom Student Notes*, which were used in the 13th century at the Esztergom cathedral school. It consisted of five parts, with the chapter on rhetoric containing Chapter IV of *Rhetorica ad Herennium*, the section on stylistic devices. At the time, much significance was attributed to grammar, which served as the basis for both dialectics and rhetoric and was recognized as the discipline of correct writing and speaking (*ars bene loquendi scribendique*). Manuscript work, not published.

Hungarian students at the time mainly studied at universities in Germany, using Latin rhetoric written by German authors, but they were also familiar with the writings of contemporary English authors. Albert Szenczi Molnár (Hungarian Reformed pastor, linguist, philosopher, psalmist, ecclesiastical writer, and literary translator) included numerous entries on rhetoric and literary theory in his rhetorical dictionaries (1604) [32]. The most significant work of our Latin linguistic literature is the book *Analyticae institutionum linguae Hungaricae* [37] by Ferenc Verseghy (Hungarian poet, polymath, literary translator, literary organizer, linguist, Pauline monk, and Catholic priest). The third part focuses on the aesthetics of our mother tongue, the purpose and sources of elegant Hungarian speech: it describes circular sentences, figures of speech, and figures of style, and then discusses the peculiarities of the Hungarian official language in an appendix.

An outstanding work from an educational point of view is the popular textbook by József Grigely (doctor of philosophy, high school principal and teacher), *Institutiones oratoriae in usum gymnasiorum Regni Hungariae at anexarum provinciarum* [17], which was used as a high school textbook in accordance with the instructions of the Ratio Educationis (two decrees of Maria Theresa's 1777 and Ferenc 1806 reform of the royal education system). Its popularity is demonstrated by the fact that it had twelve editions, the last one was published in 1847. It should be noted that this work already points towards linguistic and stylistic independence.

János Sylvester's (humanist scholar, Bible translator, founder of Hungarian essay writing) *Grammatika Hungaro-Latina* [31] is a milestone in the history of Hungarian-language education and rhetoric. It is the first Hungarian grammar book to be published in print, written by Sylvester in Latin, which sought to provide a scientific description of the Hungarian language based on ancient and humanist grammatical and rhetorical traditions. The original edition of the book was published in 1539 at the printing press in Leutschoviae (in present-day Slovakia).

The first Hungarian rhetoric, Pál Medgyesi's ecclesiastical rhetoric [22], which mainly contains his principles of preaching and rhetoric from around 1650, is an important milestone in the development of Hungarian Protestant rhetoric. Although he did not write a formal "book on rhetoric", his work is permeated by a systematic rhetorical approach and had a significant influence on school and pulpit education in the 17th-century Hungarian Reformed community.

The next significant work appeared 142 years later (in 1792) and was written by Pál Szenthe [33], the "*Mohátsi Prédikátor*" (Mohács Preacher), entitled *Magyar Oskola* (Hungarian school), with the following dedication on the title page: "For the benefit of the nation, for the advancement of its native language, and for its future refinement". It is divided into Hungarian grammar – with an emphasis on sentence structure –, Hungarian oratory, and Hungarian poetry. These two works were published primarily as school textbooks, in the form of handwritten notes or rhetorical guides for ecclesiastical purposes, and they form a bridge between the teaching rhetoric of Pál Medgyesi and the systematic rhetoric of Pál Szenthe. In Protestant colleges (e.g., Debrecen, Sárospatak, Kolozsvár, Nagyenyed), teachers and students wrote their notes on rhetoric in Latin during this period, but no remarkable were produced between the two studies mentioned above.

József Péczeli's (Reformed pastor, literary organizer, translator, and poet) series of studies, *Mindenek Gyűjtemény* (All-Hungarian Collection) [25], is a remarkable document of Hungarian literary and rhetorical culture at the end of the 18th century, containing knowledge that was modern for its time. Péczeli's multifaceted, handwritten and edited materials comprise rhetorical, literary and school collections, which played a role in education, public speaking and literary culture. In them, he emphasizes the importance of persuasion, explaining that persuasion is possible by adjusting to the needs of the audience. Printed in

1791 and published in Komárom, it can be considered to be one of the first Hungarian-language educational and literary periodicals.

Linguistic, cultural, educational, and political factors all contributed to the emergence of foreign rhetoric in Hungary in the 19th century, and it became timely for foreign (mainly German, French, and English) rhetorical works to become available in Hungarian. Among other things, there was a growing demand for the development of rhetorical skills: not only priests, but also teachers, politicians, and writers needed them. For this reason, Hugh Blair's influential book on rhetoric appeared in Hungarian translation by János Kis in 1838 [1] under the title *Blair Hugó' rhetorikai és aesthetikai leczkéi*. With some omissions and abridgements from the English version, János Kis published it in two volumes under the title *Blair Hugó' rhetorikai és aesthetikai leczkéi*. In it, the author conveyed the rhetorical understanding of the 18th-century Enlightenment, which in Blair's work was already combined with classical rhetoric (Aristotle, Cicero, Quintilian), according to which a moral-aesthetic approach is essential, i.e., the intention of a good orator is not only to persuade, but also to have a moral influence on the audience. This prestigious work is still enjoyable reading today, it is frequently quoted by rhetoricians of the reform era and later periods, is used in education, and excerpts from it are included in textbooks.

The tradition of eloquence (eloquentia) in 19th-century British rhetoric [13] also took root in Hungarian practice, appearing in the epoch-making work *Ékesszólástan* (The art of eloquence) by József Szvorényi (writer, Cistercian monk, priest, high school teacher) and Demeter Laky (educational writer and literary translator), as well as in their combined treatment of stylistics, versification, poetics, and rhetoric. The work is a rhetoric textbook, one of the most important sources for secondary schools in the second half of the 19th century. Its aim was to combine the classical rhetorical traditions with the peculiarities and examples of the Hungarian language and to place rhetoric at the service of school education and national culture. Its pedagogical significance was evident in the fact that the book was systematically adapted to the secondary school curriculum, used illustrative examples from Hungarian literature (excerpts from the works of Hungarian authors of the reform era), and encouraged speech and style exercises, and emphasized the usefulness of memorization.

After 1868, a change in terminology also took place, signaling a significant shift in approach and language use: this was accompanied by educational reforms, the reorganization of the secondary school curriculum, and the strengthening of modern pedagogical guidelines in the national language. Essentially, it was from this point on that the technical vocabulary of rhetoric as we know it today was used. At the same time, modern terminology in textbooks increasingly moved away from Greek and Latin parallels, which marked the beginning of a process of clarification and the strengthening of communication in Hungarian. The teaching material for rhetoric in Hungary gradually became more uniform, with the themes of presentation and memorization slowly falling by the wayside and being replaced by a stronger emphasis on composition, which can be attributed to the following reasons: after Hungarian became the official language (in 1844), there was a need for its regular and uniform written use so that school literacy could become a vehicle for national culture and identity. As a result, spontaneous, local variants of spoken language were pushed into the background within the framework of education by the written, standardized language.

From the second half of the 19th century, the conscious development of speaking skills appeared in teacher training institutes as part of teacher training [28]. "Beautiful speech" and "correct pronunciation" were taught as separate subjects, and the avoidance of dialects was considered important in order to strengthen the standard language. Recitation and speech contests were often organized, public speaking was practiced, and expressive, error-free speech was evaluated in teacher training exams. For example, they paid attention to intonation, pauses, and the expression of feelings during speech. The rhetorical education of

common teachers was therefore mainly based on reading, presentation and practical speech exercises.

By the 20th century, rhetoric that discussed more effective didactic views had become prevalent in education, such as the novelty of the method of imaginative illustration. The most significant authors in this field include Ferenc Acsay (Benedictine monk and teacher), Frigyes Riedl (Hungarian literary scholar and the first significant Hungarian representative of modern literary history) and Aladár Zlinszky (literary historian and aesthetician). A common feature of their works is that they extend rhetoric to the editing of prose works (e.g., drafting, analysis of parts of famous speeches and descriptions, highlighting key points) and its teaching, with a focus on historiography (see below for their introductory analysis together with a discussion of several Hungarian rhetoricians).

The institutional teaching of rhetoric was first prescribed by the 1868 law on public education for students in the sixth grade (aged 14–15) of secondary schools. A year later, the curriculum promoted by József Eötvös (Hungarian lawyer, writer, minister of religion and public education) based rhetorical knowledge on speech and comprehension exercises and composition teaching in elementary schools (for 6–12 year olds), which focused on clear Hungarian speech, the development of text comprehension and logical thinking, and the formation of good manners and community communication. The instructions accompanying the curricula emphasized that the subject should be taught through works of literature and explanatory discussion,” i.e., that the essence lies in the interpretation of the readings: the teacher helps with text comprehension, image description, story completion, and more confident argumentation by asking targeted questions.

From the beginning of the 20th century, the teaching of rhetoric shifted from theory to practice. This trend was most evident in middle schools and high schools, especially in the upper grades, as well as in teacher training colleges. In the 1924 curriculum, rhetoric remained in the 5th grade of secondary school (corresponding to today’s 9th grade, ages 14–15), but the curriculum was narrowed down to “Drawing and summarizing rhetorical lessons. An overview of prose genres.” The 1938 curriculum no longer mentions rhetoric at all.

4 The most significant Hungarian authors and works in rhetorical education in schools

After 1867, Hungarian curricula equated the theory of prose genres with rhetoric, and modern Hungarian rhetoric was born. Its novelty lay in its sophistication, scientific foundation, and use of modern terminology. This approach is most characteristically represented by the monumental work of Ferenc Acsay (a Benedictine monk and teacher), *A prózai műfajok elmélete* (The theory of prose genres) [1], which is unique in its kind; no similar summary has been written since then, in which the author attempts to create a systematic theory of prose genres: He deals with the classification of written works, description, treatise, and historiography from a school education perspective. The chapter on treatise covers logical operations such as intellectual activity and methods of generalizing knowledge, followed by the definition and classification of concepts, judgment, induction, and deduction. This chapter was followed by several later school rhetoric books. The 790-page book was only published once, probably because it was too detailed for everyday use. Textbook authors used it as a source and referred to it constantly. In addition to the classics, Acsay used the rhetoric and logic of the English author Richard Whately [38], adapting primarily his argumentation, but he also drew on Bain’s English composition and rhetoric [6] and contemporary linguistics, physical phonetics, and, above all, on the sentence-centered approach to grammar developed in the 1880s by Zsigmond Simonyi (Esperantist, linguist, university professor) (1878–1901) [30].

Frigyés Riedl's work *Rhetorika és rehtorikai olvasók* (Rhetoric and rhetorical readers) [29] was used in secondary school education for almost fifty years, with its first edition published in 1888 and its last, revised 11th edition appearing in 1933. His book introduced practical innovations based on pedagogical ideas, such as including practical exercises after each chapter in addition to the theoretical sections, applying rhetorical studies not only to oral but also to written genres (which was a new approach at the time), and practicing style imitation, i.e., imitating the style of famous writers or orators for learning purposes. The illustrative tables served to make the material transparent and systematic, while the selected extensive world literature and Hungarian reading material served to introduce and analyze rhetorical concepts, figures of speech, and stylistic devices. The preface also contains some advice worth considering: "Let all your words be for the sake of meaning, and meaning for the sake of truth" (1888), the books illustrated in Figure 1.

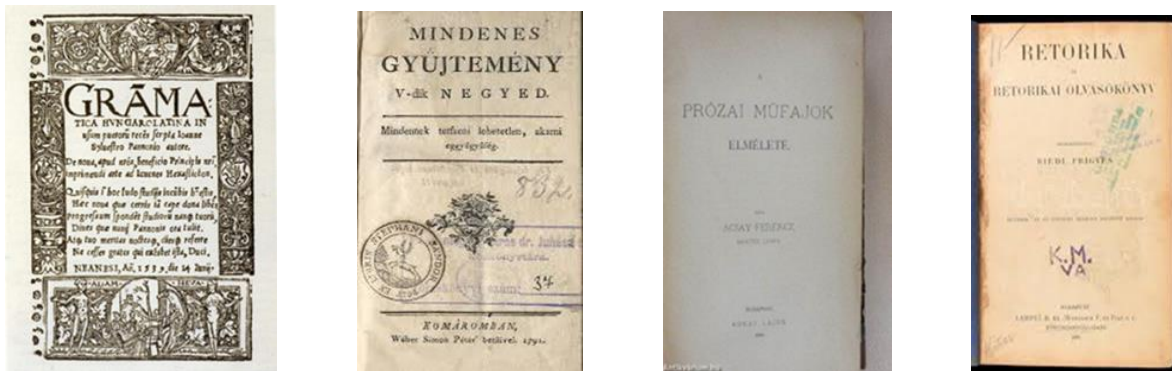


Figure 1. The covers of the original editions of four Hungarian rhetorical books: *Grammatica Hungaro-Latina*, *Mindenes Gyűjtemény* [All-Hungarian Collection], *A prózai műfajok elmélete* [Theory of prose genres], *Rhetorika és rehtorikai olvasókönyv* [Rhetoric and rhetorical reading book]

In 1907, Aladár Zlinszky published his work entitled *A prózai műfajok mintái és törvényei. Középiscolák számára* (Examples and rules of prose genres. For secondary schools) [39]. It is a voluminous book of 272 pages, divided into reading material (excerpts from novels and novellas by contemporary Hungarian writers, informative and philosophical prose, speech and letter genres, descriptions, characterizations) and a theoretical section (in which he seeks logical and linguistic regularities highlighting clear, concise expression and examining prose genres from a functional rather than a literary-historical perspective). In the preface, he explains: "Our instructions set out the path for reform when rhetoric was replaced by the theory of prose genres in general, rather than the theory of oratory. In this respect, our schoolbooks were ahead of similar foreign rhetoric textbooks, which still deal only with the theory of oratory" (first, unnumbered page).

In 1901, László Négyesy published *Rhetorika. Prózai olvasmányok. Elmélet. Gymnasiumok és reáliskolák V. osztálya* (Rhetoric. Prose readings. Theory. For gymnasium and real schools, grade V) (corresponding to today's 9th grade) [23], which, like Aladár Zlinszky's work, follows an inductive approach and is divided into readings and theoretical sections. It is a clear, readable textbook. Its structure is as follows: 1. narratives (including historiography), 2. treatises, 3. letters (a large collection of Roman and Hungarian letters), 4. instructional works (e.g., moral instructions), 5. speeches (notable among these are the *Memorial Speech for Ferenc Kazinczy* by Kölcsey and *Kossuth's speech on June 11, 1848*). The chapter entitled "speech" interprets the concept of speech as a live, persuasive work, and the types of speech can be political, judicial, ecclesiastical, academic, and occasional, including memorial speeches, moral exhortations, letters, greetings, and funeral speeches.

Also worth mentioning is László Névy and his textbook entitled *Rhetorika. A prózai műfajok elmélete. Középiscolák és magánosok használatára* (Rhetoric: A theory of prose

genres. For use in secondary schools and private education) [24] in 1878. This work marks the beginning of a series of rhetoric textbooks linked to the school curriculum. It launched a view linked to legal requirements that placed native language, practical composition, and the development of thinking at the center of rhetorical education. The structure of his textbook follows the definitions of rhetoric, according to which there are broader and narrower interpretations: in a broader sense, it refers to the theory of all prose genres in which eloquence (*eloquentia*) is dominant, while in a narrower sense, it refers to the theory of public speaking. In addition to a detailed discussion of theoretical knowledge, he also covers the practice of successful speech delivery. This part is referred to in the classical division of rhetoric as *actio* or *pronuntiatio*, i.e., the manner of delivery, the tone of voice, body language, and the use of voice are essential. In his opinion, prolonged practice enables the speaker to deliver a speech that is well received even if based on a “bare” outline or even improvised. Through practice, they become familiar with their own oratorical abilities and become familiar with the field and medium of speech. They feel at home in front of an audience, and thus the pithy expression of their thoughts always comes to their lips at the right moment. This is the ideal goal that students of oratory must keep in mind.

Szónoklattan (The Art of Public Speaking) is a work by István Verbőczy [36], a Cistercian teacher in Pécs. It was published in Budapest in 1870 and 1878 for use in schools and for private study. The textbook is actually the work of József Szvorényi (linguist, literary historian, educator, Cistercian monk, and high school principal) which went through seven editions, *Ékesszólástan, vezérletül a remekírók fejtegetése s a szép írásművek kidolgozásában* (1851–1877) (The art of eloquent speech, guided by the explanations of great writers and the development of beautiful writings) [34]. It provides the most comprehensive description of the types of arguments. His book is not aligned with the curriculum, and he adds stylistics to rhetoric.

Imre Góbi, a high school teacher, only discusses historiography and oratory under the title of rhetoric in his 1888 work *Rhetorika. A gymnasiumok VI. osztálya számára* (Rhetoric. For sixth-year secondary school students) (aged 14–15) [15]. It can be described as a practical textbook, as its theoretical material is short and concise, but it contains a wealth of examples so that teachers can introduce the theory to their students through readings. The author provided outlines of the readings to facilitate preparation at home. Illustrated with specific examples: 1. the task is to give a simple and clear account of real or imaginary events in the narrative; 2. the aim is for the listener (reader) to understand what happened, when, where, how, and why; 3. requirements: adherence to chronological order, conciseness and clarity, avoidance of unnecessary details; 4. the expected style is clear, direct, and natural. The textbook also provided an example task for the outline: *Write a narrative about how you helped a stranger on the street! (The aim of the task: moral message, adherence to chronological order, clear expression).*

János Dengi’s textbook, *Rhetorika. A prózai műfajok elmélete. Olvasmányokkal s ezek magyarázatával középiskolai használatra* (Rhetoric. A theory of prose genres. With readings and explanations for secondary school use) [14], can be classified among the works of Nény and Góbi, and includes detailed analysis of rhetoric, historiography, and oratory. It also includes expository prose in the core material. His collection of examples includes readings that reflect the understanding of speech in his own time. In line with the general understanding of his time, Dengi regarded speech not only as a means of communication, but also as one of the foundations of moral and intellectual education. In his opinion, good speech is not only beautiful, but also meaningful, clear, and serves the truth. He marks the parts of the speeches, the place of the arguments, and provides explanations and structural outlines.

Kálmán Pintér’s work provides an overview of prose in general, the concept of rhetoric, rhetorical style, and prose genres, entitled *Rhetorika és olvasókönyv. Iskolai használatra*

(Rhetoric and reader. For school use) [27]. The work has a clear, outline structure, offering short, concise definitions, logically structured subchapters, and material organized into points. This form makes it easier for students to understand and memorize, and for teachers to divide the material into sections.

5 Conclusions

Rhetorical education in Hungary has been continuous since the 13th century and has undergone a process of development. The first rhetoric textbooks, written in Latin and later in Hungarian, are based on the works of ancient and Anglo-Saxon theorists from both a theoretical and historical perspective, as well as on knowledge of grammar. Institutional teaching was introduced in Hungary by educational legislation in the second half of the 17th century. The writing and use of textbooks in secondary education began intensively in the second half of the 19th century. Many professional linguists became textbook authors as practicing teachers, who, in addition to teaching the theoretical knowledge of domestic rhetoric education, continuously enriched its methodological palette in line with the spirit of the times and with the aim of developing student's oral and written expression. They contributed greatly to the modern interpretation of rhetoric as a metalanguage (discourse about discourse).

During the 19th and 20th centuries, the teaching of rhetoric underwent significant changes in Hungary. By the end of the 19th century, classical, school-based rhetoric training had gradually faded into the background, increasingly replaced by grammar and literature education. At the turn of the century, rhetoric no longer appeared as a separate subject, but was scattered throughout other subjects.

The history of rhetorical education in Hungary from the 13th to the 19th century clearly reflects the changes in Hungarian educational policy, society, and culture. The gradual transformation and "disappearance" of traditional rhetorical education did not mean that the demand for rhetorical skills disappeared, but only that they came to the fore in a different form and in a new context.

References

- [1] ACSAY, F.: *A prózai művek elmélete* [Theory of prose works], Laiszky János Könyvnyomdája, Esztergom, 1889.
- [2] ACZÉL, P.: *A retorika tanítása* [The teaching of rhetoric], *Anyanyelv-pedagógia*, IX. évfolyam, 2. szám, 2016, <https://doi.org/10.21030/anyp.2016.2.1>.
- [3] ADAMIK, T. (ed.): *Retorikai lexikon* [Rhetorical lexicon], Kalligram, Bratislava, 2010.
- [4] ADAMIK, T., A. A. JÁSZÓ, and P. ACZÉL: *Retorika* [Rhetoric], Osiris, Budapest, 2005.
- [5] ARISTOTELES: *Rétorika* [Rhetoric], Telosz, Budapest, 1999.
- [6] BAIN, A.: *English composition and rhetoric*, Appleton, New York, 1867.
- [7] BERLIN, J., A.: *Writing Instruction in nineteenth-century American colleges*, Southern Illinois University Press, 1984.
- [8] BLAIR, H.: *Lectures on rhetoric and belles lettres*, A. Strahan, T. Cadell, and W. Creech, Edinburgh, 1787.
- [9] BOOTH, W. C.: Foreword. Seriously considering rhetoric education. In: J. PETRAGLIA, B. DEEPIKA (eds.): *The realms of rhetoric: The prospects for rhetoric education*, State University of New York Press. Albany, New York, 2003.
- [10] CAMPBELL, G.: *The philosophy of rhetoric*, Harper, New York, 1868.
- [11] CORNIFICIUS: *A C. Herenniusnak ajánlott retorika* [The rhetoric proposed to C. Herennius], Akadémiai Kiadó, Budapest, 1987.
- [12] CRYSTAL, D.: *A nyelv enciklopédiája*. [The encyclopedia of language], Osiris, Budapest, 2003.

- [13] CONLEY, T. M. *Rhetoric in the European tradition*, The University of Chicago Press, Chicago, 1990.
- [14] DENGL, J.: *Rhetorika. A prózai műfajok elmélete* [Rhetoric. Theory of prose genres], Eggenberger, Budapest, 1890.
- [15] GÓBI, I.: *Retorika. A gymnásiumok VI. osztálya számára* [Rhetoric. For class VI in grammar schools], Franklin Társulat, Budapest, 1888.
- [16] GÓSY, M.: *Pszicholingvisztika* [Psycholinguistics], Osiris, Budapest, 2005.
- [17] GRIGELY, J.: *Institutiones oratoriae in usum gymnasiorum Regni Hungariae at anexarum provinciarum*, Editio Nova, Budae, 1823.
- [18] HAWKINS, E. W.: *Awareness of language: an introduction*, Cambridge University Press, Cambridge, 1984.
- [19] KIS J.: *Blair Hugó' rhetorikai és esztetikai leckéi. Némely kihagyásokkal és rövidítésekkel angolból* [Hugo Blair rhetorical and aesthetic tasks. With some omissions and abbreviations from English], Magyar Királyi Egyetem, Buda, 1838.
- [20] LEVELT, W. J. M.: *Speaking. From intention to articulation*, A Bradford Book, Cambridge, Massachusetts, 1989.
- [21] LIBÁRDI, P.: *A mentális lexikon több szempontú vizsgálata* [A multifaceted examination of the mental lexicon], Eötvös Kiadó, Budapest, 2019.
- [22] MEDGYESI, P.: *Doce nos orare, quin et praedicare*, Klösz Jakab, Bártfa, 1690.
- [23] NÉGYESY, L.: *Retorika a Gymnasium V. osztálya számára* [Rhetoric for class V of the gymnasium], Lampel, Budapest, 1901.
- [24] NÉVY, L.: *Rhetorika. A prózai műfajok elmélete* [Rhetoric. Theory of prose genres], Eggenberger, Budapest, 1878.
- [25] PÉCZELI, J.: *Mindenes Gyűjtemény* [All-Hungarian Collection], Wéber Simon Péter, Komárom, 1791.
- [26] PINKER, S.: *A nyelvi ösztön* [The language instinct], Typotex, Budapest, 2006.
- [27] PINTÉR, K.: *Rhetorika és olvasókönyv* [Rhetoric and reading book], Hornyányszky, Budapest, 1895.
- [28] PÓLYA, I.: *A tanítóképzés és a pedagógiai gyakorlat Magyarországon a 19. század második felében* [Teacher training and pedagogical practice in Hungary in the second half of the 19th century], Akadémiai Kiadó, Budapest, 2001.
- [29] RIEDL, F.: *Rhetorika és rehtorikai olvasókönyv* [Rhetoric and rhetorical reading book], Lampel Róbert (Wodianer F. és Fiai) Cs. és Kir. Udvari Könyvkereskedés, Budapest, 1888.
- [30] SIMONYI, ZS.: *A magyar kötőszók, egyúttal az összetett mondat elmélete* [Hungarian conjunctions and the theory of the compound sentence], Magyar Tudományos Akadémia, Budapest, 1881.
- [31] SYLVESTER, J.: *Grammatika Hungaro-Latina*, Sárvár-Újsziget, 1539.
- [32] SZENCZI MOLNÁR, A.: *Dictionarium Ungarico-Latinum*. Hutterus, Noribergae, 1604.
- [33] SZENTHE, P.: *Magyar Oskola* [Hungarian School], Pest, 1792.
- [34] SZVORÉNYI, J., D. LAKY.: *Ékesszólástan* [Eloquence], Heckenast Gusztáv, Pest, 1851/1870.
- [35] QUINTILIANUS, M. F.: *Institutio oratoria*, Kalligram, Budapest, 2009.
- [36] VERBŐCZY, I.: *Szónoklattan. Iskolai és magánhasználatra* [Rhetoric. For school and private use], Aigner L., Budapest, 1870/1878.
- [37] VERSEGHY, F.: *Analyticae Institutionum Linguae Hungaricae*, Typis Regiae Universitatis Hungaricae, Budae, 1817.
- [38] WHATELY, R.: *Rhetoric*, Charles Griffin and Co., London, 1872.
- [39] ZLINSZKY, A.: *Rhetorika. A prózai műfajok mintái és törvényei* [Rhetoric. Patterns and laws of prose genres], Lampel Róbert (Wodianer F. és fiai), Budapest, 1907.

BETWEEN SOUND AND SCRIPT: HENRY SWEET'S PHONETICS IN LATE VICTORIAN BRITAIN

*Qi Hong*¹

¹*University of Toronto, Canada
qi.hong@mail.utoronto.ca*

Abstract: This paper reexamines Henry Sweet's "living philology" within the broader context of late Victorian language sciences, approaching his work through the lens of the history of science. This paper argues that Sweet's work exemplifies how language cannot be fully abstracted from its material supports, nor reduced to a purely physiological or physical phenomenon. The first section focuses on Sweet's proposal of phonetics being the first requisite for studying the living philology. It examines the philologist's attention to the synchronicity and the diachronic history of language. With the "living philology" explained, the second section dwells on Sweet's insistence on the primacy of the observer. In this section, I contextualize Sweet's techniques of observing in relation to the epistemic virtue of mechanical objectivity. Sweet's negativism against using inscriptive instruments in phonetic experiments indicates an attempt to redefine objectivity. The third section launches an analysis of Sweet's engagement with the technological condition of the production of linguistic knowledge. Although Sweet was explicitly critical of instrumental phonetics, his development of a universal phonetic transcription system nonetheless relied on print technologies, revealing a more ambivalent relationship to technological mediation than his rhetoric might suggest. The transformation of linguistic knowledge, it concludes, entails not only theoretical innovation but also the reconfiguration of representational systems, technological apparatuses, and their interrelations.

1 Introduction

In 1884, British philologist Henry Sweet wrote, "[T]he living form of every language should be made the foundation of its study... for the spoken is the only form of the language which is regular and definitely limited in the range of its grammar and vocabulary" [1]. Sweet's view that the study of language should begin with its spoken form was not monstrously unique among European physiologists, physicists, and phoneticians in the late nineteenth century. Instead, already in the eighteenth century, the understanding of human speech as an acoustic phenomenon was already flourishing, mostly fashioned by European architects, musicians, and engineers [2]. In the nineteenth century, emerging fields of physiology and acoustic science further transformed the study of speech with not only new devices for observation and measurement but also new analytical tools such as the analysis of waveforms and signals [3]. Clearly, physiological and physical research on human speech was hardly anything revolutionary in late nineteenth-century Europe.

However, what makes Henry Sweet and his proposal for philologists to focus on the "living form of language" noticeable to historians of language and historians of science is his own positionality as a British philologist. While Sweet was calling for a turn from studying grammar and vocabulary to studying speech phenomena, Sweet was also aware of the "antagonism between practical linguistic phonetics and instrumental phonetics," and publicly stood against the latter [4]. Sweet's knowledge of and criticism of phoneticians who rely on mechanical devices was by no means a total rejection of phonetic studies. Instead, formed in the late nine-

teenth century, his proposal that “phonetics should be the first requisite to study language” [5] was made at a time when disciplinary boundaries between linguistic sciences were being formed. As a philologist interested in studying speech, Sweet became a thorny figure for scholars to categorize—not only for phoneticians who wanted to write their own history of the discipline [6] but also for historians of science who wanted to understand epistemological transformations through tracing the formation of modern scientific disciplines [7].

This paper, written from the perspective of a historian of language sciences, historicizes Henry Sweet and his “living philology” in late Victorian Britain. By taking Henry Sweet as an epitome of the paradigm shift in language sciences, the first two sections of this paper situate Sweet in relation to his predecessors and contemporaries. The first section focuses on Sweet’s proposal of phonetics being the first requisite for studying the living philology. It examines the philologist’s attention to the synchronicity and the diachronic history of language. With the “living philology” explained, the second section dwells on Sweet’s insistence on the primacy of the observer. In this section, I contextualize Sweet’s techniques of observing in relation to the nineteenth century’s “epistemic virtue of mechanical objectivity,” identified by Lorraine Daston and Peter Galison in their work *Objectivity* [24]. The persona of an ideal phonetician leads to a critical reflection on Sweet’s negativism against instrumental phonetics, which is not merely an irrational rejection of modern experimental science. The third section launches an analysis of Sweet’s engagement with the technological condition of the production of linguistic knowledge. Although Sweet was explicitly critical of instrumental phonetics, his development of a universal phonetic transcription system nonetheless relied on print technologies, revealing a more ambivalent relationship to technological mediation than his rhetoric might suggest. The transformation of linguistic knowledge, as this paper argues, entails not only theoretical innovation but also the reconfiguration of representational systems, technological apparatuses, and their interrelations.

2 The Living Philology

In 1875, Henry Sweet read a paper to the London Philological Society. This paper, titled “Words, Logic, and Grammar,” begins with a statement on the philologist’s responsibility to observe the living phenomena of language:

“Philologists forget... that the history of language is not one of decay only, but also of reconstruction and regeneration. ... [These processes] have further the great advantage of being perfectly accessible to the observer. ... But before history must come a knowledge of what now exists. We must learn to observe things as they are, without regard to their origin, just as a zoologist must learn to describe accurately a horse, or any other animal” [9].

About a year before he read this paper, Sweet had finished his research on the sound-changes of Middle English by comparing it with Scandinavian languages. This work, *History of English Sounds* (1874), along with Sweet’s previous publications on Old Icelandic grammar and vocabulary, is very much still based on textual materials to trace changes in linguistic sound since antiquity. But Sweet did not conform to the general expectation of a philologist who only builds one’s theories of a linguistic past solely on a textual corpus. In another 1874 report on the recent trends in Teutonic languages philology, Sweet exhibited a clear favor for the physiology of vocalization over antiquarian philology, positively hoping for a “general revolution in philological method,” which he saw in the German trend [10]. Sweet was particularly interested in works by German physicians Ernst Wilhelm von Brücke (1819–1892) and Carl Ludwig Merkel (1812–1876). He also showed some sufficient knowledge of physiology as he stated that Merkel’s work was more advanced than Brücke’s, although Sweet did not offer any rationale behind his opinion [10].

By the time Sweet reoriented the focus of philology to “what now exists” in “Words, Logic, and Grammar,” phonetics had yet to become what it is now understood as: a branch of linguistics that studies human production and perception of linguistic sounds. Contrastingly, before Sweet, a “phonetician” in mid-to-late Victorian scholarly discourse referred to a scholar working on the categorization of speech sounds using notational symbols, who oftentimes was the inventor of the system of symbols. Two predecessors whom Sweet frequently referred to in his works are Alexander J. Ellis (1814–1890) and Alexander Melville Bell (1819–1905), both of whom Sweet praised for being attentive to details of speech sounds. Both Ellis’s and Bell’s works share two preoccupations: on the one hand, an accurate classification of vowels and consonants based on pronunciation; on the other hand, a writing system that efficiently represents acoustic differences in speech [11]. Ellis and Bell relied heavily on acoustic and physiological knowledge. For instance, in *The Essentials of Phonetics* (1847), Ellis quoted Robert Willis (1800–1875) at length to define phonetics as a branch of acoustic studies, while he also borrowed from Karl Moritz Rapp’s (1803–1883) work on the physiology of speech and language [12]. Bell, in his seminal work *Visible Speech* (1867), based his classification of vowels on the modification of the guttural passage, the buccal cavities, and the lips (also known as “rounding”), as well as the position of the tongue during vocalization [13]. Nevertheless, it remains unclear to what extent the early British phoneticians were fully acquainted with the experimental methods and instruments used by acousticians and physiologists, and whether early phoneticians like Ellis were able to apply these approaches and equipment to their own analyses of speech [14]. Hence, it must be noted that, despite the reliance on acoustics and physiology, early British phonetic research was still largely about bridging the gap between the writing system and the spoken language. For Ellis and Bell (and to some extent, also Sweet), the analysis of human speech always rests upon a practical concern: How should one efficiently transcribe what one hears? New symbolic systems—alphabets in these cases—were invented as a result, be it Ellis and Pitman’s phonotypy, or Pitman’s own phonography, or Bell’s vowel chart. Overall, studies of speech sound in mid-to-late Victorian Britain, as J. A. Kemp perceptively concludes, developed within the context of language studies, with strong emphasis on phonetic scripts, language pedagogy, and philology [15].

Indeed, despite his frequent references to early phoneticians’ works, Sweet did not fully agree with them on the goal of having an effective shorthand system—not to mention his bitterness toward Pitman phonography, referring to it as “the popular ‘Pitfall’ system” [16]. Unsatisfied by the unhealthy favoring of brevity over accuracy in existing shorthand systems, Sweet invented his own system, the “Current Shorthand,” in 1892. While Sweet’s rationale for improving sound notation systems will be discussed in detail in the following section, it is important for now to emphasize that, for Sweet, the necessity of having an efficient transcribing system goes far beyond the everyday need for speech transcription. Shorthand systems were invented to register linguistic changes by dissecting speech sounds into phonetic units, which can offer more accurate data on changes in pronunciation than the written language. In other words, whilst the living form of language refers to speech and pronunciation, the ultimate mission of the living philology is still to understand how languages have changed throughout the historical course. But this change, for Sweet, is not merely diachronic; it also has a synchronic dimension. As he elaborated in the quoted paragraph at the beginning of this section, the history of language is not just about decay but also about “reconstruction and regeneration” [9]. The living form of language, that is, various pronunciations of the same word across different regions, is a crucial clue to capturing linguistic changes. According to the same passage, the synchronic aspect of language conditions access to the knowledge of its history. Under this logic, there would be no a priori structure of language that transcends the current time-space. The observable data, or in Sweet’s terms, the “actual facts of language,” are what deserve the philologist’s attention [17].

In conclusion, unlike the practical phoneticians, Sweet believed that the shorthand system serves a scholarly pursuit. The living philology values the synchronic aspect of language as much as it is concerned with the shared origin of language families [18]. In this sense, the living philology prioritizes the spoken sound as the primary source of analysis. Hence, instead of relying solely on the written texts, living philology entails an efficient transcription system for obtaining data. How would the living philology differ from experimental phonetics? Why would Sweet believe that the symbolic system of transcription is better than the inscriptive and recording machines—such as the kymograph and the X-ray—widely used by the experimentalists? By tracing Sweet’s rationale behind his invention of the “Current Shorthand,” the following section discusses the role of the observer and the techniques of observing in Sweet’s philological-phonetic research.

3 Techniques of the Observer

In 1877, Sweet took presidency of the Philological Society. In the same year, one of Sweet’s most famous works *A Handbook of Phonetics* (hereafter, the handbook) came out. In his presidential address as well as in the preface of the handbook, Sweet accounted for a history of practical phonetics in Britain beginning from Max Müller to A. J. Ellis, and then to A. M. Bell. With their interest in the spoken languages shared, Sweet acknowledged that he inherited Ellis’s notation system at large, and Bell’s “clear and concise” terminologies. He also highlighted his alterations of former phonetic research—for instance, he proved that Bell’s “wide” and “primary” distinction can also be applied to consonants—and also stated that he corrected some misleading conclusions [19].

Nonetheless, Sweet carefully distanced himself from Ellis, Bell, and other phoneticians and grammarians. Most noticeably, he not only emphasized the practicality of phonetics not only in foreign language pedagogy but also implied that the acquisition of a foreign language or dialect should be a prerequisite for phoneticians. Sweet even identified the native language as one of the phonetician’s biggest burdens in acquiring accurate description of linguistic sounds, stating:

“A great deal depends on the character of the native language, the learner [of phonetics] naturally grafting the peculiarities of his own language on his pronunciation of foreign ones, as when an Englishman diphthongises the long vowels in French and German; and, again, finding those sounds difficult which do not occur, or have no analogues, in his own pronunciation. It is a great mistake to suppose that any one nation has a special gift for acquiring sounds or foreign languages generally. Each nation has its special defects or advantages” [20].

Talking about foreign language acquisition, however, Sweet ultimately aimed to train sensitive and acute observers of unwritten dialects and spoken languages [21]. The educated observer, in the ideal case, is both a native of the dialect to be investigated and a trained phonetician who could dissect the individual unit of their speech accurately. Moreover, the spoken dialects and the literary language have a dialectical relationship, in the sense that “literary English will soon efface all the still existing remains of our dialects” [21]. The ideal observer, therefore, bears two primary burdens. On the one hand, one has to know their native pronunciation patterns, so that one can register phonetic differences accurately [20]. On the other hand, the observer must be able to perform the “unrestrained, natural dialect” as to be free from the “forced, unidiomatic sentences, full of words and forms taken from the literary language” [21]. This makes it extremely hard for one to become a qualified observer, as Sweet also acknowledged, because even with thorough training and practice, it is still impossible to avoid making any mistakes in hearing and interpretation.

Given his instances on the precision of observation, it would be reasonable for Sweet to switch to experimental phonetics, with which the phenomenon of speech has been equated to

acoustic and/or aerodynamic mechanisms, conveniently becoming observable through the mechanical instrument. Yet, instead of replacing the naturally inaccurate human observer with the more objective inscriptive device, Sweet resorted to several non-experimental approaches to make up for the unavoidable errors. Immediately, Sweet proposed the method of “direct questioning,” which asks the observer to examine educated dialect speakers who can discern literary expression and natural colloquial usage. The difficulty of this method is to find such speakers, as to master the dialect, one needs to pay more efforts than merely acquiring several sounds or intonations, or some isolated expressions [22].

Hence, the knowledge of practical phonetics comes back to Sweet as an important technique for the phonetician to acquire the features of any spoken language. The notation system is a helpful tool not only for transcribing speech sounds but also for understanding one’s own native speech. Such a system must also be applicable to different systems of pronunciation. Based on the roman letters, Sweet created the “Romic alphabet,” which seeks to represent phonetic variations without exhausting the innumerable possible combinations of segments into syllables. Consisting of two sub-systems, “Broad Romic” and “Narrow Romic,” the Romic alphabet is adaptive to suit the requirement of each specific language. The Broad Romic system marks all necessary distinctions of sound in one single language in the simplest and most direct manner. It is used to register distinctions of sound which correspond to distinctions of meaning in the target language [23]. Unlike the broad system, the narrow system does not register the phonemic unit. Instead, symbols in the narrow system reference the physicality and acoustics of the spoken sounds. It aims to characterize each sound as minutely accurate as possible, thus operating on a comparatively more universal level of representation. For instance, the short and narrow /i/ and the short and wide /i/ (italicized letter “i”) signify different meanings in Danish and Icelandic, while in French, there is no such distinction, and in British English, the difference between narrowness and wideness only occur between the short /i/ and the longer /i:/ . In this sense, using the broad system for the English language, the letter “i” would be enough for representing both the short, wide /i/ and the short, narrow /i/ [23]. Similarly, to transcribe British English with the broad system, one does not need to italicize the letter “i” in the /i:/ , as wideness comes naturally when the vowel is prolonged. But if one wants to compare the British sound with that of Danish and/or Icelandic, the narrow system is of great use, because the formal difference between “i” and “i” must be registered as they refer to different degrees of roundness and positions of the tongue. The narrow system, in this sense, records the physiological and articulative aspects of speech that transcends linguistic barriers, while the broad system notes the phonological difference within each specific language.

Counterintuitively, Sweet believed that the Romic Alphabet is more accurate than the instruments for measurement and inscription. Sweet’s negativism against instrumental phonetics is well known, but the rationale behind is likely dismissed as being subjective or unscientific. This current paper, however, aims to contextualize Sweet’s persistence on the human observer in phonetic research in what Lorraine Daston and Peter Galison call “the epistemic virtue of objectivity,” which emerged in mid-nineteenth century [24]. According to Daston and Galison, “mechanical objectivity” originally characterizes the repression of the artist-author of the atlas from intervening in the making of the drawings and diagrams in mid-to-late nineteenth century scientific publications [25]. It indicates the emergence of a scientist persona, who would rely on mechanical devices to obtain data, and would refrain from imposing their own interpretation extensively. Indeed, Daston and Galison never suggest that mechanical objectivity would completely replace, and be replaced by, previous and succeeding epistemic virtues. They also never assert that scientists in the era of mechanical objectivity must subjugate themselves to machines and protocols. This leaves historians with the possibility to ponder whether the scientific persona characterized by Daston and Galison is, once and for all, absolute [26].

Contextualized in the era of mechanical objectivity, Sweet's negativism against instrumental phonetics leads to a critical re-examination of mechanical objectivity as something more than a truism. The problem with experimental phonetics, for Sweet, is twofold. First, machines used for collecting phonetic data, such as the phonautograph and the phonograph, are also sources of errors and mistakes. By mediating the process of data collection, these machines prevent the phonetician from directly accessing the facts of languages. Second, the experimental method is but only a help. The phonetician is the final arbiter to note down what is being heard, and one needs to have a trained ear to be qualified for the role of the observer. Human perception was, in other words, not erased by, nor subordinated to, the rapid advancement of transcriptive and inscriptive tools. Sweet's view neither assumes the phonetic structure as a natural ideal (as he frequently emphasized that the living language is always changing), nor does it subjugate human subjectivity under machines. Instead, it is perceptive to foresee the epistemic value of "trained judgement," that is, the following stage after mechanical objectivity, which flourished in the mid-twentieth century, according to Daston and Galison. In other words, the era of mechanical objectivity contains within itself its own antithesis. Henry Sweet's techniques of observing, from learning a foreign language to direct questioning, and to the Romic Alphabet, can thus be read as a critique of mechanical objectivity.

4 The Technological *a priori*

The invention of the Romic Alphabet is more than an intellectual endeavor. The Roman alphabet failed to meet Sweet's demands because of its arbitrary relation to the sound it represents. A relatively better representational system, according to Sweet, is Bell's visible speech. In this alphabet each letter symbolizes the action of the vocal organs by which it is formed, according to certain fixed principles [27]. Yet, other than its incapacity in accurately represent speech sounds across languages, another important reason that kept Sweet from endorsing Bell's visible speech is its complexity. In spite of its imperfections, the Roman alphabet remains to be the optimal choice because it has long been in practical use. Producing new types is costly; also, these types further disturb and complicate the existing founts. But the existing orthography can always be reformed without expense or disturbance of the existing machinery of the printing-offices [28].

Sweet, however, was not the only person who was aware of the practical concern of making a new notation system. A decade earlier, Alexander J. Ellis already wrote, "It appeared to me desirable to have an alphabet consisting entirely of those types which we may expect to find in every printing office and hence consisting only of Roman and Italic letters [...]" [29]. With the "Palaeograph" notation system, Ellis attempted to document every possible speech sound with a unique symbol. Most strikingly, he also self-printed his own *Essentials of Phonetics* (originally titled, ESENSALZ OV FONETICS) at his own printing office [30]. This endeavor, for Sweet, seems to be too much trouble. Acknowledging Ellis's attention to the practical matter and that the Roman alphabet "can only be a temporary compromise," Sweet maintained that what is at stake is to find signs for fundamental distinctions, and the exceptional ones can easily be supplied by simple description [31].

Simplicity and efficiency were key principles for Sweet to propose a new notation system based on the existing Roman alphabet. In *A Handbook of Phonetics*, he proposed four principles to make up for the "imperfections of the Roman alphabet," namely, by casting a few new letter types; by employing diacritics, such as grave and acute accents; by employing digraphs, such as "th" and "kh" by employing turned letters, italics and capitals [32]. For instance, to print "ə", a nineteenth-century typesetter would simply put a linotype "e" upside down on a composition stick. The Roman alphabet, in this sense, is not just a collection of letters. It is both material and symbolic, consisting of a limited number of signs and metal types. But the relation between the metal type and the letter sign ceased to be exclusive in the case of the

Romic alphabet. One type could represent multiple symbols. In other words, while each phonetic symbol represents one phonetic unit (or a set of similar units, or allophones), the process of materializing the representation requires the heavy mediation of a “technological a priori,” where the instrument is no longer just a means to an end, but a fundamental condition of representation [33].

The materiality of the Roman alphabet embodied by printing later became the backbone of the Roman alphabet as the primary source for making notation systems. In 1886, Paul Passey founded “Dhi Fonètik Títcer’z Asóciécon” (F.T.A.) in Paris, France, which later on turned into l’Association Phonétique Internationale, or the International Association of Phonetics, which is known for inventing a universal system of notation, the International Phonetic Alphabet (IPA) [34]. In the first issue of the association’s journal, *Le Maître Phonétique*, was printed in the phonetic alphabet adopted from Pitman and Ellis’s 1847 alphabet [35]. Modifications were made, such as “c” and “j” were used in place of “f” and “z” because the French printing house which typeset the journal did not have a full set of fonts for all letters [36]. But still, the Roman alphabet sustained its dominant position because of its connection to printing technology, in particular the types. Or in other words, for any phonetic alphabet based on the roman letters, adding new symbols to the existing alphabet is only possible when the printing technology permits. As Robert Brain points out, “Linguistic values... had to be arbitrary in order to remain robust, but natural in order to remain valid. ... [S]tandards were semiotic and semiotics were standards” [37]. The basic principle of aligning all symbols according to the Roman alphabet has not been changed throughout the century after. In 1989, the IPA’s principle states, “The non-roman symbols of the IPA have, as far as possible, been made to harmonize with the roman letters” [38].

5 Conclusion

Despite the popular experiments on speech reproduction and the pioneering works in phonetic notations, in late Victorian Britain, phonetics was yet to become a well-defined discipline. Also, it was not widely accepted by the British academia of language sciences, where philology was dominant. Henry Sweet, in spite of being elected as the president of several associations, such as the Philological Society in 1878 and the IPA from 1897 to 1912, never got a chair in any British university. “In future ages, it will seem incredible that in the latter half of the nineteenth century there was not a single authorized teacher of phonetics ... in England, the natural home of phonetics [...]” [39]. Sweet criticized his concurrent British philologists’ lack of interest in the study of the “living language,” while he also rejected the instrumental approach that would reduce speech sounds to mere physical or physiological phenomena. Sweet’s insistence on the skillful observer foreshadows a scientific persona that was yet to come, while his major invention, the Romic Alphabet, hints at a lurking contradiction between phonetics and phonology. While experimental phonetics soon became a major trend worldwide in the 20th century, the Roman alphabet as the primary notational system soon also encountered various challenges and failures in orthographic systems where letterpress printing was alien. As a conclusion, the history of linguistics is not only a history of experimental science but also one of symbolic systems and information technologies.

References

- [1] SWEET, H.: *The practical study of language*. In: H. C. WYLD (ed.): *Collected Papers of Henry Sweet*, p. 35, THE CLARENDON PRESS, 1913. Italicized by this author.
- [2] GILLIN, E. J.: *Science and Sound in Nineteenth-Century Britain: Sounds Experimental and Entertaining*, p. 20, 2023.
- [3] ASHBY, M.: *Experimental phonetics in Britain, 1890-1940*, p. 117, 2016.
- [4] ASHBY, M.: *Experimental phonetics in Britain, 1890-1940*, pp. 146–147, 2016.
- [5] SWEET, H.: *Presidential Address (1877)*. In: H. C. WYLD (ed.): *Collected Papers of Henry*

- Sweet, p. 85, THE CLARENDON PRESS, 1913.
- [6] OHALA, J. J.: *The integration of phonetics and phonology. Proceedings of the XIIIth International Congress of Phonetic Sciences, Aix-en-Provence. Vol. 1*, pp. 1-16, 1991.
This work was later adapted by John J. Ohala into “Phonetics and Phonology then, and then, and now.” In this work, Ohala categorizes Sweet’s works, as well as other IPA-related phonetician’s works as “taxonomic phonetics,” despite that he also noticed the duality of Jean-Pierre Rousselot and Paul Passy.
- [7] CHANG H.: *Inventing Temperature: Measurement and Scientific Progress*. OXFORD UNIVERSITY PRESS, 2007.
For a more elaborated account on the difference between history & philosophy of science (HPS) and scientific research (including its disciplinary history), I find Hasok Chang’s concept of “complementary science” very useful.
- [8] BRAIN, R.: *Standards and Semiotics*. In: T. LENOIR (ed.): *Inscribing Science: Scientific Texts and the Materiality of Communication*, pp. 250–252, STANFORD UNIVERSITY PRESS, 1998.
- [9] SWEET, H.: *Words, logic, and grammar*. In: *Transactions of the Philological Society, 1875–6*, pp. 470–503, reprinted in H. C. WYLD (ed.): *Collected Writings*. Emphases are mine.
- [10] SWEET, H.: *On Germanic and Scandinavian (1874)*. In: H. C. WYLD (ed.): *Collected Papers of Henry Sweet*, p. 75, THE CLARENDON PRESS, 1913.
Sweet also shows a clear favor of Merkel over Brucke, although both physiologists used diagrams to represent the vocal tract in the production of vowel sounds.
- [11] ELLIS, A. J.: *The Essentials of Phonetics*, p. 23, 1848.
In 1845, Alexander J. Ellis (1814–1890) published *The Alphabet of Nature*. Two years later, the second edition of *The Alphabet of Nature* came out under the new title, *The Essentials of Phonetics*. In both works, Ellis explained a shorthand for the English language and a “theory for a universal alphabet.”
- [12] WILLIS, R.: *On the Vowel Sounds, and on Reed Organ-pipes*. In: *Transactions of the Cambridge Philosophical Society*, pp. 1–38, 1829.
Robert Willis (1800–1875), engineer and college fellow at Cambridge, built a machine that imitates vowel sounds using pipes with an adjustable plunger. For more details, see Gillin, *Science and Sound in Nineteenth-Century Britain*, p. 67.
- [13] BELL, A. M.: *Visible Speech*, 1868.
For a more comprehensive review on the development of phonetics in mid-victorian Britain, see Michael Ashby, *Experimental Phonetics in Britain, 1890–1940* (2016); and J. A. Kemp, “The development of phonetics from the late 18th to the late 19th century,” pp. 1475–1479.
- [14] ASHBY, M.: *Experimental phonetics in Britain, 1890-1940*, p. 118, 2016.
On Ellis’s usage of the phonautographic diagrams, Michael Ashby acutely points out that it is “unclear whether this means that Ellis had access to a device and was able to make tracings, or simply that Wheatstone (whom Ellis referenced extensively) had been sent many examples from other workers.”
- [15] KEMP, J. A.: *The development of phonetics from the late 18th to the late 19th century, and The history and development of a universal phonetic alphabet in the 19th century: from the beginnings to the establishment of the IPA*. In: *History of the Language Sciences: An International Handbook on the Evolution of the Study of Language from the Beginnings to the Present*. pp. 1572–1580, 2000.
- [16] SWEET, H.: *A Manual of Current Shorthand Orthographic and Phonetic*, p. 16, 1892.
- [17] SWEET, H.: *Linguistic affinities (1900)*. In: H. C. WYLD (ed.): *Collected Papers of Henry Sweet*, p. 56, THE CLARENDON PRESS, 1913.
- [18] SWEET, H.: *Preface*. In: *A Handbook of Phonetics: Including a Popular Exposition of the*

Principles of Spelling Reform, p. 13, 1877.

- [19] SWEET, H.: *Preface*. In: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, pp. 15–16, 1877.
- [20] SWEET, H.: *English and Germanic Philology (1878)*. In: H. C. WYLD (ed.): *Collected Papers of Henry Sweet*, p. 130, THE CLARENDON PRESS, 1913.
Sweet saw a dialectical relation between spoken dialects and literary languages. For instance, contemporary Italian dialects are not “evolved” (or “decayed”) forms of an old spoken Italian, but are “direct decedents of literary Latin,” which “swallowed up all the old Italic dialects long before the rise of the Romance languages.”
- [21] SWEET, H.: *English and Germanic Philology (1878)*. In: H. C. WYLD (ed.): *Collected Papers of Henry Sweet*, p. 128, THE CLARENDON PRESS, 1913.
- [22] SWEET, H.: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, pp. 103–104, 1877.
- [23] SWEET, H.: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, pp. 106–107, 1877.
- [24] DASTON, L. AND GALISON, P.: *Objectivity*. ZONE BOOKS, p. 17, 2010.
- [25] DASTON, L. AND GALISON, P.: *Objectivity*. ZONE BOOKS, p. 121, 2010.
- [26] YEANG, C.: *Mechanization of hearing in Chao Yuan Ren’s dialect research: Senses, Objectivity, and Observation, Chinese Annals of History of Science and Technology*, pp. 94–117, 2019.
Historians of science and technology have shown that the subjugation of the self in scientific research is not always the case. For instance, Chen-Panng Yeang shows that, in the 1920s and 30s, human data collectors, such as Chinese linguist Yuan Ren Chao, continued to play a central part in phonetic research in China. Chao’s usage of his “tone spectral measurers” required delicate operations as well as interventions by the human data collector with a “musical ear,” as it requires the observer’s input to generate results. In this sense, Henry Sweet was not the only one who characterized an alternative scientific persona.
- [27] SWEET, H.: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, p. 171, 1877.
- [28] SWEET, H.: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, p. 174, 1877.
- [29] ELLIS, A. J.: *On Palaeotype; or, the representation of spoken sounds, for philological purposes, by means and the ancient types, Transactions of the Philological Society*, p. 3, 1867.
- [30] ELLIS, A. J.: *Essentials of Phonetics*, 1848.
- [31] SWEET, H.: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, p. 102, 1877.
- [32] SWEET, H.: *A Handbook of Phonetics: Including a Popular Exposition of the Principles of Spelling Reform*, p. 101, 1877.
- [33] KITTLER, F. A.: *Gramophone, Film, Typewriter*. STANFORD UNIVERSITY PRESS, 1999.
- [34] MACMAHON, M. K. C.: *The International Phonetic Association: the first 100 years*. In: *Journal of the International Phonetic Association*, p. 30, 1986.
- [35] N. A.: *Hû Wî Ar*. In: *Le Maître Phonétique*. 1886.
- [36] MACMAHON, M. K. C.: *The International Phonetic Association: the first 100 years*. In: *Journal of the International Phonetic Association*, p. 36, 1986.
- [37] BRAIN, R.: *Standards and Semiotics*. In: T. LENOIR (ed.): *Inscribing Science: Scientific Texts and the Materiality of Communication*, pp. 283–284, STANFORD UNIVERSITY PRESS, 1998.
- [38] INTERNATIONAL PHONETIC ASSOCIATION: *The principles of the International Phonetic Association*. In: *Handbook of the International Phonetic Association: A Guide to the*

Use of the International Phonetic Alphabet. CAMBRIDGE UNIVERSITY PRESS, p. 205, 2005. Emphasis mine.

[39] SWEET, H.: *Practical study of language (1884)*. In: H. C. WYLD (ed.): *Collected Papers of Henry Sweet* pp. 50, 1913.

OTTO BREMER - PHONETICIAN, FIELD LINGUIST AND FOUNDER OF THE PHONETIC COLLECTION AT HALLE UNIVERSITY

Sven Grawunder^{1,2}, Ursula Hirschfeld¹

¹Martin Luther University Halle-Wittenberg, Germany, ²Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
grawunder@sprechwiss.uni-halle.de, hirschfeld@sprechwiss.uni-halle.de

Abstract: The dialect researcher and phonetician Otto Bremer founded the Phonetic Collection in 1910. He had worked at Halle University since 1888. He became known as a vivid teacher, enthusiastic researcher, and field worker of German dialects and the Frisian language. In 1937, the Phonetic Collection was united with the collections of speech science, continued in 1947 under the name “Institut für Sprechkunde und Phonetische Sammlung” and is now part of the Department of Speech Science and Phonetics. It consists of historical equipment, some of which was already used by Bremer, which was used to record, preserve, and examine spoken language, and the *Schallarchiv*, a sound archive, which also contains original recordings made by Bremer and now contains more than 12,000 recordings from the various fields of speech science.

1 Otto Bremer

Otto Bremer - philologist, field linguist, phonetician, writer, language activist, and founder of the Phonetic Collection at Halle University was a prolific academic writer, a vivid university teacher, and an active academic. And, he has over the past two decades become himself a subject of research, especially with regard to the history of the two sister disciplines - phonetics and speech science - at Halle university.

1.1 Short Biography

Otto Bremer was born in Stralsund on November 22, 1862, as son of a bookseller. After attending grammar school (*Gymnasium*) in Stralsund (1871–1881), Bremer studied German philology and comparative linguistics at the universities of Berlin, Heidelberg and Leipzig from 1881 to 1886. In Leipzig he studied with Friedrich Zarnckes, August Leskien, Karl Brugmann, and Eduard Sievers.

In Leipzig, Bremer received his doctorate in 1885 with a thesis on the phonetic history of the Germanic development of Indo-European ē [1]. In 1888, he completed his habilitation at the *Vereinigte Friedrichs-Universität Halle-Wittenberg* (the united university Halle-Wittenberg was a merger of the two universities Leucorea and Friedrichs-Universität in 1817) with a thesis on the ‘Amringian-Föhringian dialect’ [2], i.e. the dialect of the islands Amrum and Föhr, and Bremer was awarded the *Venia Legendi* for “German Language and Literature”.

Otto Bremer worked as a private lecturer at the German Department for over eleven years before he was awarded the title of professor in 1899, although this was not linked to any appointment. It was not until 1904, after repeated applications from the faculty, that the responsible minister appointed him as a tenured associate professor for the subject of “Germanic Philology”.



(a) Otto Bremer, 1894



(b) Otto Bremer, 1931

Figure 1 – source: www.catalogus-professorum-halensis.de

Between 1910 and 1918, he was first reappointed to the subject of “Phonetics and General Linguistics”, then in 1919 to “Phonetics, General Linguistics and German Dialect Studies”.

Only after his 65th birthday in 1928 - but before the end of his period of service (end of the winter semester 1927/28) on March 13, 1928 - was Bremer appointed full professor; as a result of this appointment, he was given emeritus status on April 1, 1928 instead of retirement, which later partially protected him. With the entry into force of the Nuremberg laws, Otto Bremer saw his teaching license revoked as a “non-Aryan professor” with effect from 31 December 1935. He died of cancer on August 8, 1936 and was buried in the municipal Gertrauden Cemetery in Halle (Saale).

2 Research

2.1 Research on dialects

Otto Bremer earned a reputation through numerous dialectological works, a still remarkable phonetic precision, and exemplary empirical studies on the pronunciation realities of various varieties.

Bremer had already devoted himself to the study of German dialects with his postdoctoral thesis (Habilitationsschrift)[2]. He made significant contributions to German dialectology, the study of Frisian, phonetics, Germanic history, and theoretical questions of language change. He developed the theoretical and methodological foundations for the study of local grammars in the late 19th century and implemented them in a model fashion through the series he edited and partly authored, the “Collection of Short Grammars of German Dialects” (Sammlung kurzer Grammatiken deutscher Mundarten), for example ...

- Bremer, O. (1893) *Deutsche Phonetik. (Sammlung kurzer Grammatiken deutscher Mundarten Vol. I)*. Breitkopf & Härtel, Leipzig. (reprinted Vaduz 1984)
- Bremer, O. (1895). *Beiträge zur Geographie der deutschen Mundarten in Form einer Kritik*

von Wenkers Sprachatlas des Deutschen Reichs (Vol. III). Breitkopf & Härtel.

- Gebhardt, A., & Bremer, O. (1907). *Grammatik der Nürnberger Mundart (Vol. VII)*. Breitkopf & Härtel. (reprinted 1968)
- Gerbet, E. (1908). *Grammatik der Mundart des Vogtlandes. Lautlehre (O. Bremer, Ed.; Vol. VIII)*. Breitkopf & Härtel.
- Kürsten, O., & Bremer, O. (1910). *Lautlehre der Mundart von Buttstedt bei Weimar (Vol. IX)*. Breitkopf & Härtel.

Aware of both the difficulty and the necessity of graphically representing spoken dialect, he created a standardized phonetic script based on his phonetic studies. Bremer had already applied the approach realized in his Collection of Short Grammars of German Dialects - describing and explaining the current sound conditions of dialects in their historical development - in his classification and mapping of German dialects, which was included in the 14th edition of the Brockhaus Encyclopedia in 1892 [3]. Together with Otto Behagel's work in 1891 [4] this is considered to be one of the "first scientifically based maps of the German dialects" [5].

2.2 Discourse with Wenker

Some remarkable early anecdote is Bremer's critique of Georg Wenker's Sprachatlas in Marburg, which was compiled of 30.000 questionnaires of school teachers across the new German Reich after 1871. Wenker was the appointed director of the research institute for German language in Marburg; first publications came out in 1881. The third volume of Otto Bremer's collection series (1895) was explicitly dedicated to a critique on Wenker, mainly with the focus on the lack of precision in those laymen's orthographic transcriptions. This critique was swiftly returned by Wenker and Wrede with a volume subtitled "Poetry and Truth" [6] the same year and rebutted by Bremer [7] the following year, finally re-analysed by others [8].

2.3 General Phonetics

As mentioned above, Bremer was aware that writing can only incompletely and distortedly represent spoken language. His own suggestions of a phonetic transcription system, first published in his "Deutsche Phonetik" [9], sought to accommodate those shortcomings (s. Figure 2) [cf. 10].

As Hoffmann [12] points out, the precise sagittal sections accompanied by palatograms in the "Deutsche Phonetik" can be counted as milestones in the depiction of speech sounds. With the invention of the phonograph in 1877, it became possible for the first time to record sound and reproduce it later. Bremer recognized its value for his own research and teaching and persistently advocated for a collection of phonetic-acoustic materials and equipment in Halle. In 1910, he received financial support from the university and founded the Phonetic Collection.

2.4 Wangeroog Frisian

Otto Bremer first traveled to Wangeroog and Neu-Wangeroog in 1898 to study this variety of Frisian. Many visits followed. In 1924, he recorded examples of the dialect, which was already threatened with extinction at the time, on phonograph cylinders, some of which have been preserved in the Phonetic Collection (see below), as has the phonograph he used. Bremer's last wish was for these to be transferred to shellac records, which his successor Richard Wittsack carried out. Based on the phonograms, Wittsack [13, 14] also published a short book on

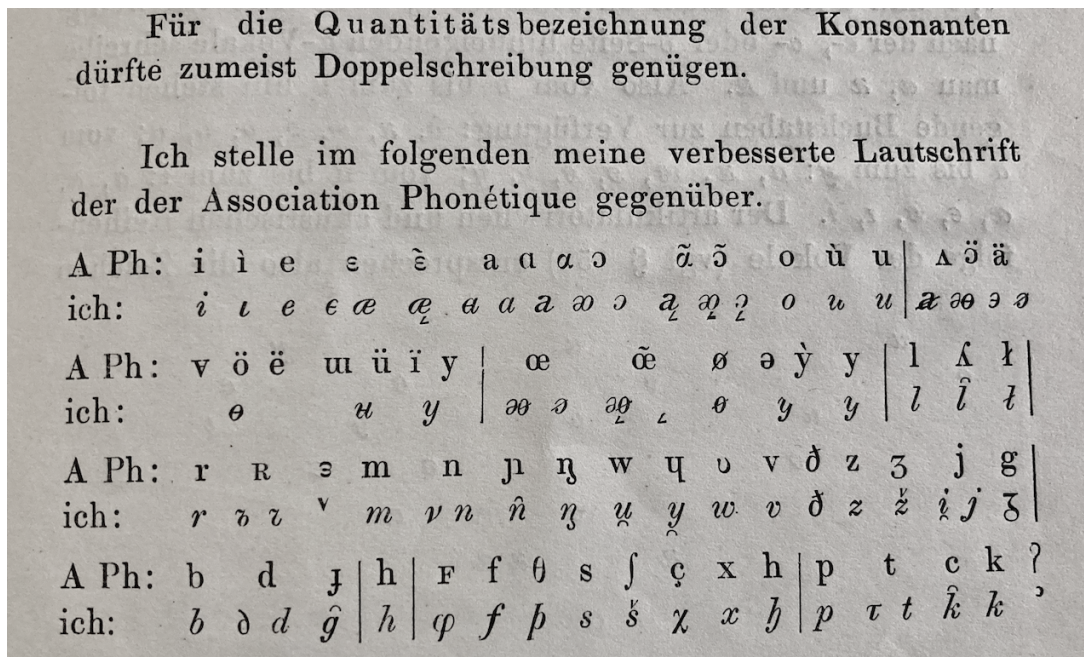


Figure 2 – Bremer’s suggestions vis-à-vis the *Association Phonétique*, [11, p.11]

Wangeroog Frisian in 1938 with contributions from Dietrich Gerhardt for the linguistic analysis and Edgar Fuhrhop for the description of the transfer process to shellac at Telefunken (Berlin).

2.5 Other academic activities

Otto Bremer was honored for his scientific achievements by several learned societies and associations: in 1888, he became a corresponding member of the *Friesch Genootschap van geschied-oudheid-en taalkunde* in Leeuwarden; in 1892, a corresponding member of the *Leeuwarden Selskip for Fryske taal- en skriftenkennisse*; and in 1903, an honorary member of the Association for Transylvanian Regional Studies (*Verein für siebenbürgische Landeskunde*).

In addition to research and university teaching, Bremer was also active in practical language work on the ground. He rendered outstanding services to the Halle dialect, and until 1925, he served for more than 15 years as chairman of the Halle branch of the German Language Association (*Deutscher Sprachverein*).

2.6 Research about Bremer

Although Bremer’s collection remained an active part of the university’s holdings, particularly in the field of speech studies, it was the more recent active engagement of historians [15] investigating the history of the excluded Jewish professors that brought Otto Bremer himself back into the focus of the discipline and its history [16]. There are questions that remain currently unanswered, e.g. with regard to Bremer’s somewhat unusual academic career. There is his very late appointment as adjunct professor and his even later appointment as full professor. Solms [17] points out that until his habilitation work at 25 years, everything looks fairly common, compared to his colleagues in the field of German language studies. However, it appears Bremer’s perhaps somewhat eccentric personality, his new kind of thinking with regard to fieldwork and his mere ‘linguistic’ profile of publications or other disputes with the local more traditional scholars of German philology [cf. 17] which caused a difficult stand and thus a lack of support? Another curious story is that of Bremer’s Mission in Frisland in 1920, when he was expelled from North Frisland by the local authorities for interfering with the local referendum about independence or belonging to Germany or Denmark [cf. 10].

3 Bremer's heritage

3.1 Phonetic Collection

In 1910, Bremer received financial support from the university and founded the Phonetic Collection. Under this name, he brought together various devices, such as the Kymographion – his teacher Sievers in Leipzig had used this device beforehand –, the Edison Phonograph (wax cylinder apparatus) and the Parlograph (dictation machine with wax cylinder) (s. Figure 3a), with which he was able to record and analyze numerous dialects. He had people from different German-speaking regions speak test sentences in their dialect, also known as Wenker sentences, and archived the recordings [18]. Some of these early recordings are still preserved in the Schallarchiv today, including recordings of Wangeroog Frisian (s.below).

In 1922, the collection was recognized by the state by ministerial decree and was given the status of an independent university institution.

In 1925, the speech scientist Richard Wittsack (1887-1952) from Halle founded and established a further speech collection.

In 1937, the collections of Otto Bremer and Richard Wittsack were merged and Richard Wittsack took over the management of the expanded Phonetic Collection until 1952.

In 1947, the Phonetic Collection and the Institute's own holdings became the "Institute for Speech Studies and Phonetic Collection" (*Institut für Sprechkunde und Phonetische Sammlung*).



(a) Parlograph (1912) by Carl Lindström, AG Berlin, photo: Maike Glöckner
(b) Telefunken decelith recorder / player (1937), photo: Maike Glöckner

Figure 3 – Recording devices: two generations

By 1926, the collection was already well equipped. According to a report by Bremer in the university chronicle, it contained a kymographion, a Flatau's laryngoscope, a laryngeal mirror after Gutzmann, a series of larynx and half-head models, the speaking organs in alcohol, a skull, five anatomical wall charts, an overtone chart, Gutzmann's apparatus for demonstrating his theory of vowels, a Gutzmann bottle for demonstrating the diaphragm and lungs, an air pump, a manometer, a metronome, a stopwatch, several mirrors, tuning forks, tuning pipes, a phonograph with many cylinders, three speaking machines with various metal and wooden funnels, sound boxes and listening tubes and around 170 records and cylinders for many languages, especially for German dialects and for exemplary performance, as well as a Pathé and a Plantor recording apparatus, plus various of his own cylinders and recordings. [...] a laryngeal mirror (autophonescope) according to Panconcelli-Calzia and a hinged nose model" (Chronicle of the Vereinigte Friedrichs-Universität Halle-Wittenberg, for the period 1916 - 26, quoted from



(a) bottom of storing tube for the wax cylinder with Wangeroog Frisian, original handwriting by O. Bremer; photo: Maike Glöckner



(b) English speakers from New York and New Jersey, 'natural talk', 'recorded with long funnel'; photo: Maike Glöckner

Figure 4 – Examples of the storing tubes in the collection

Lemmer [19, 193]).

In addition to the old Kymographion from Bremer's collection, which used a blade of grass as a pen, two other, more modern devices from the Leipzig companies Zimmermann and Diel were acquired in the 1940s by Richard Wittsack and in the 1960s by Eduard Kurka.

Over the following decades, the Phonetic Collection was an ever-present part of linguistic research and teaching in Halle. The sound archive was also continuously supplemented by collecting and archiving contemporary poetry, speeches, therapy courses, radio recordings, etc.

However, the storage conditions of the collection were more or less poor for a long time, depending on the respective locations of the institute: in boxes and shelves in cellars and store-rooms [20]. It was not until 2015, with the move to the Steintorcampus, that suitable storage and exhibition facilities were available. Eventually, it was also possible to provide workstations for assistants.

The Department of Speech Science and Phonetics continues to collect recordings and equipment, as well as to prepare historical audio and video material for research, archive it on current data carriers, and record it in databases.

Central pieces of Bremer's collection are the Kymographion, the Edison phonograph, the Parlograph. However, also later acquisitions like the *Telefunken Reiseapparatur* for mobile recordings on decelith sound foils (s. Figure 3b) are rare and valuable pieces of the collection.

Work is currently underway to incorporate the collection items into the university's central catalogue of Custody (<https://sammlungsportal2.kustodie.uni-halle.de/>).

3.2 Wax cylinders of Bremer

42 of the 121 cylinders in the sound archive can be clearly attributed to Otto Bremer, including all 13 recordings of the Wangeroog Frisian dialect. These contain his handwritten notes at the bottom of the storing tube (s. Figure 4a):

- Title of the recording ("Wangerooger Friesisch"),
- date and the location of the recording ("Wangerooge, 24.9.24"),
- the funnel used ("long funnel"),
- recording speed ("7 revolutions less than maximum speed"),
- personal details about the speaker ("The last person on the island who still speaks the language fluently and purely."),
- information on the speaking style ("spoken very slowly"),
- content ("natural narrative from her life").

Curiously, there are also about 5 cylinders and storing tubes which remain telling about

Bremer's documentation of his daughter's language acquisition at the age of 4 to 5 years. Also in this aspect, Bremer seems to be influenced by the neogrammarians. Other cylinders reveal recordings of foreign languages and foreign accents (s. Figure 4b), like French, English, "Andalusian", Lithuanian, or Ukrainian spoken by non-native and native speakers.

3.3 Sound Archive

The department of speech science and phonetics maintains, among others, a focus on applied phonetics in the context of teaching German as a foreign language and continued to collect those kinds of recordings. Further poetic texts, monologues i.e. political speeches but also examples and documentations of language, speech and voice disorders and their therapy were collected. The *Schallarchiv* (sound archive), which Bremer had started, comprises today around 10,000 individual recordings, including 1250 shellac and wax records, 110 cylinders, around 220 sound foils, 1,400 studio tapes and other reel-to-reel tapes recorded at the Institute, as well as numerous other media such as audio cassettes, DAT cassettes, minidisks and CDs. The collection not only documents historical and current speech science research, it also serves as a source for further research and is of general interest. The department's current goal is to make the collection fully accessible and available to the scientific community.

The *Schallarchiv*'s website provides some examples:

<https://www.schallarchiv.uni-halle.de/>

<https://www.schallarchiv.uni-halle.de/phonetik/>

<https://www.schallarchiv.uni-halle.de/rhetorik/>

<https://www.schallarchiv.uni-halle.de/sprechkunst/>

<https://www.schallarchiv.uni-halle.de/stimmstoerungen/>

<https://www.schallarchiv.uni-halle.de/geraetesammlung/>

<https://www.schallarchiv.uni-halle.de/archiv-uebersicht-schallplatten/>

4 Acknowledgments

We would like to thank our long-time employee and technician Mr. Peter Müller for his help in providing the photographic material and for his great personal commitment to the preservation and further growth and prosperity of the Phonetic Collection. Further, we are in debt to Mrs. Waltraud Rösel, Otto Bremer's granddaughter, for her generous moral and financial support of the many works contributing to the Phonetic Collection.

References

- [1] BREMER, O.: *Germanisches ē 1. Teil. Die lautgesetzliche entwicklung des idg. ē in den ältesten germanischen Sprachen*. Karras, Halle (Saale), 1885.
- [2] BREMER, O.: *Einleitung zu einer amringisch-föhringischen Sprachlehre*. Soltau, Norden [u.a.], 1888.
- [3] BREMER, O.: *Deutsche Mundarten*. In *Brockhaus' Konversationslexikon*, vol. 5, pp. 27–35. Brockhaus, Leipzig, 14 edn., 1892.
- [4] BEHAGHEL, O.: *Geschichte der deutschen Sprache*. Strassburg K.J. Trübner, 1891.
- [5] WIESINGER, P.: *Probleme der Dialektgliederung des Deutschen*. *Zeitschrift für Dialektologie und Linguistik*, 49(2), pp. 145–168, 1982.

- [6] WENKER, G. and F. WREDE: *Der Sprachatlas des deutschen Reichs: Dichtung und Wahrheit*. Elwert, Marburg, 1895.
- [7] BREMER, O.: *ZUR KRITIK DES SPRACHATLAS. Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)*, 1896(21), 1896. doi:10.1515/bgsl.1896.1896.21.27.
- [8] HEUSLER, A.: *Bremer Beiträge zur Geographie der deutschen Mundarten in Form einer Kritik von Wenkers Sprachatlas des deutschen Reichs. Indogermanische Forschungen*, 8(s1), pp. 96–99, 1897. doi:10.1515/if-1897-0165.
- [9] BREMER, O.: *Deutsche Phonetik*. Sammlung kurzer Grammatiken deutscher Mundarten 1. Breitkopf & Härtel, Leipzig, 1893.
- [10] HÜTTNER, A.: *Otto Bremer als Phonetiker – Ein Beitrag zur sprechwissenschaftlichen Fachgeschichte*. In *Otto Bremer - Wegbereiter der sprechwissenschaftlichen Phonetik an der Universität Halle*, no. SSP 5 in Schriften zur Sprechwissenschaft und Phonetik, pp. 61–145. Frank und Timme GmbH, Berlin, 2016.
- [11] BREMER, O.: *Zur Lautschrift*. Sammlung Kurzer Grammatiken Deutscher Mundarten, Anhang Zu Band 1. Breitkopf & Härtel, 1898.
- [12] HOFFMANN, R.: *Early activities in the field of X-ray imaging in experimental phonetics and speech training*. In *Proc. 20th Int. Congr. of Phonetic Sciences (ICPhS 2023)*, pp. 7–11. 2023.
- [13] GERHARDT, D. and E. FUHRHOP: *Wangerooger Friesisch Studienergebnisse an Hand der Wangerooger Phonogramme der Phonetischen Sammlung der Universität Halle*. Phonetische Sammlung der Martin-Luther-Universität Halle-Wittenberg 1. Phonet. Sammlung der Martin-Luther-Univ., Halle (Saale), 1938.
- [14] MITZKA, W.: *Review of Wangerooger Friesisch. Studienergebnisse an Hand der Wangerooger Phonogramme der Phonetischen Sammlung der Universität Halle*. *Zeitschrift für Mundartforschung*, 15(2), pp. 128–128, 1939.
- [15] STENGEL, F.: *Ausgeschlossen zum Gedenken an die 1933 - 1945 entlassenen Hochschullehrer der Martin-Luther-Universität Halle-Wittenberg*. Univ.-Verl. Halle-Wittenberg, Halle an der Saale, 1. Aufl. edn., 2013.
- [16] HIRSCHFELD, U., H.-J. SOLMS, and E. STOCK: *Otto Bremer - Wegbereiter der sprechwissenschaftlichen Phonetik an der Universität Halle*. No. SSP 5 in Schriften zur Sprechwissenschaft und Phonetik. Frank und Timme GmbH, 2016.
- [17] SOLMS, H.-J.: *Wer wider den Stachel löckt – Zum beruflichen Lebensweg Otto Bremers*. In *Otto Bremer - Wegbereiter der sprechwissenschaftlichen Phonetik an der Universität Halle*, Schriften zur Sprechwissenschaft und Phonetik, pp. 47–59. Frank und Timme GmbH, 2016.
- [18] STOCK, E. and U. HOLLMACH: *Die phonetische Sammlung*. In MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG and R.-T. SPELER (eds.), *300 Jahre Universität Halle, 1694-1994: Schätze Aus Den Sammlungen Und Kabinetten*, pp. 249–258. Edition STEKOFOTO, Halle an der Saale, 1994.

- [19] LEMMER, M.: *Zur Geschichte der Sprechkunde an der Universität Halle. Wissenschaftliche Zeitschrift. Gesellschafts- und sprachwissenschaftliche Reihe*, 6(2), pp. 189–196, 1957.
- [20] HÜTTNER, A.: *Zur Entwicklung der sprechwissenschaftlichen Phonetik an der Universität Halle (Saale) bis 1961: Mit einem Ausblick auf die weitere Entwicklung des Fachs*. No. SPP 15 in *Schriften zur Sprechwissenschaft und Phonetik*. Frank & Timme, Berlin, 2020.

EXPERIMENTAL PHONETICS ON DISPLAY: HERITAGE, SCIENCE, AND MEMORY AT THE UNIVERSITY OF COIMBRA FACULTY OF LETTERS

Quintino Lopes¹, Ângela Salgueiro¹

¹*University of Évora, Institute of Contemporary History / IN2PAST, Portugal*
quintinolopes1@gmail.com

Abstract: In 2025, the PHONLAB R&D project team will create a permanent exhibition at the University of Coimbra dedicated to the phonetician Armando de Lacerda (1902-1984) and the Experimental Phonetics Laboratory of the Faculty of Letters (1936-1979). The exhibition is the result of a joint effort by the Institute of Contemporary History, the Foundation for Science and Technology, the Rectory of the University of Coimbra, the Science Museum of the University of Coimbra, and the Faculty of Letters of Coimbra. This paper reports on the curatorial process and the scientific content that will be on display, organised into the following four sections: 1) Lacerda's pioneering work, 2) the work of the Coimbra phonetics laboratory, 3) the laboratory's collaborators, and 4) the creation of a sound archive of Portuguese regional varieties. The realisation of the exhibition presupposes a profound change in how the University of Coimbra perceives and promotes itself, as well as its scientific spaces and personnel. Due to open in November 2025 in the space formerly occupied by the phonetics laboratory at the Faculty of Letters — in a corridor to be named the Armando de Lacerda Wing — the exhibition will be permanent and bilingual, and will be open to the public free of charge. It is hoped that it will promote public interest in the history of experimental phonetics.

1 Introduction

In recent years, there has been a proliferation of scientific work on the phonetician Armando de Lacerda (1902–1984), and the Experimental Phonetics Laboratory at the University of Coimbra, which was founded by the Institute for High Culture and directed by Lacerda between 1936 and 1979 [1], [2], [3]. Interest in the subject has increased substantially within the international scientific community — *Lacerda 120. 5th International Workshop on the History of Speech Communication Research* (Porto, 2022) [4], and *Armando de Lacerda: A Pioneer of Experimental Phonetics* (Royal Swedish Academy of Sciences, 2023) [5], and with public funding from the PHONLAB project – *Phonetics Laboratory: Coimbra - Harvard. Rethinking 20th-century scientific centres and peripheries* (Foundation for Science and Technology) [6].

The PHONLAB project offers a fresh perspective on the Portuguese scientific landscape by combining research into the history of phonetics in Portugal with an in-depth examination of the Coimbra Experimental Phonetics Laboratory and the importance of preserving the institution's scientific and material heritage. The project forms part of the national endeavour to preserve and enhance scientific and technical heritage, in accordance with the directives and recommendations provided by national and international organisations, including the UNESCO National Commission [7]. At the same time, the project ensures the continuation of work begun in 2017 to recover the historical memory of a 'forgotten' space and scientific discipline, overlooked by the

University of Coimbra. For example, in 2019 António Almeida (*1946), a disciple of Armando de Lacerda, tried to persuade the Faculty of Letters at the University of Coimbra to create a room in homage to his master — an initiative that was not realised.

Thus, one of the main objectives of the PHONLAB project is to organise a permanent exhibition at the Faculty of Letters of Coimbra – in the space formerly occupied by the Experimental Phonetics Laboratory – dedicated to the history of phonetics in Portugal, to the work of the Experimental Phonetics Laboratory and to the phonetician who ran it. This exhibition will take an approach that challenges the dichotomy of centres versus peripheries, based on the methodology of global history [1], [3].

The exhibition has been designed to achieve several of the objectives set out in the project's application process. These objectives include preserving and showcasing the material and immaterial heritage linked to the Experimental Phonetics Laboratory, which was previously spread across various reserves at the University of Coimbra. Another objective is to revive the memory of this scientific space and its protagonists by musealising instruments, archive documents, and oral testimonies into museum exhibits. This will be achieved by critically re-examining an often-overlooked scientific history [8].

This initiative is pertinent for a number of reasons. Firstly, it contributes to the conservation of scientific collections and makes them available to the public. On the other hand, it fulfils the commitment of higher education institutions to engage more with the community, upholding the principles of open science – namely citizen science, and enhancing the societal impact of science by promoting access to heritage as an educational, identity-forming and cultural resource. Furthermore, it contributes to the recognition and affirmation of scientific culture as an integral part of the Portuguese cultural history.

This paper aims to present the exhibition to be inaugurated in 2025 at the University of Coimbra Faculty of Letters. It will describe the exhibition's structure, conceptual framework, and content, including the scientific research conducted by the laboratory, as well as the teaching, dissemination, and university outreach activities promoted by the laboratory's director and its specialised researchers.

The curatorial process sought to integrate the materiality of the scientific objects with the scientific and institutional narratives that give them meaning. These narratives are organised into four sections, each of which gives its name to one of the points presented in this paper: 1) Armando de Lacerda's pioneering work on coarticulation and chromography; 2) the University of Coimbra's Experimental Phonetics Laboratory; 3) the international scientific attraction of the Coimbra Phonetics Laboratory; and 4) the Sound Archive of Portuguese Regional Speeches. These sections provide cross-references between the materiality of the objects and instruments used in phonetic studies, and oral testimonies from former collaborators and students, as well as documents — textual, iconographic and audio — that have been recovered and preserved in the Experimental Phonetics Laboratory archive and other public and private archives. These include the Paulo de Lacerda Family Archive [8].

The work in progress has endeavoured to involve all of the project's partner institutions: the Institute of Contemporary History; the University of Évora; the University of Coimbra, through its Rectorate, Faculty of Letters, and Science Museum; the University of Porto; the University of Stockholm; and the University of Macau. The aim is to guarantee inter-institutional cooperation, which is fundamental to the international dissemination of the exhibition. The curatorship involves specialists from

different scientific fields, and the content is made accessible to diverse audiences through inclusive mediation strategies. The exhibition, which will be known as the 'Armando de Lacerda Wing', also has the support of the International Phonetic Association.

In a Europe where exhibitions on the history of experimental phonetics are scarce, the magnificent HAPS collection at the Technical University of Dresden is a notable exception [9], [10], [11]. With the inauguration of this permanent exhibition at the University of Coimbra, we also aim to encourage reflection on the history of experimental phonetics and stimulate public interest in speech studies.

2 Armando de Lacerda's pioneering work. Coarticulation and chromography

The first section of the exhibition showcases Armando de Lacerda's contributions to the development of experimental phonetics. As mentioned in [2: 5], Armando de Lacerda, founder and director of the Laboratory of Experimental Phonetics at the Faculty of Letters of the University of Coimbra, was a pioneer in the field of twentieth-century speech sciences. In 1932, while specialising in Germany in the new field of experimental phonetics, Lacerda invented the polychromograph, a type of inkjet oscillograph which overcame many of the limitations of kymography, the antiquated method which remained in use at experimental phonetics laboratories up until the mid-20th century. In 1933, in collaboration with the German phonetician Paul Menzerath (1883–1954), he developed the concept of coarticulation, which has played a major role in Phonetic theory ever since. Along with the recognition of Lacerda's pioneering role in the development of new inventions, this innovative work in the field of the study of the coarticulation and segmentation of speech sounds led to his achieving lasting international prestige.

The exhibition of the book *Koartikulation, Steuerung und Lautabgrenzung*, with which Menzerath and Lacerda formulated in 1933 the key concept of coarticulation (the influence of contiguous speech sounds on each other) is accompanied by a photograph of the instrument that allowed the authors to obtain the published conclusions: Lacerda's Oral Recording Labiograph, coupled to Menzerath's Kymograph, which was installed at the University of Bonn.

Regarding chromography, the polychromograph is documented, and photographs show its application at the University of Bonn's Institute of Phonetics in 1932 and 1933. Of particular note are the rare materials on display in this sector, such as photographs of Paul Menzerath and the Institute of Phonetics in Bonn, which was destroyed during the Second World War. On display are also chromograms and technical drawings of chromographs, as well as the programme for the 1st International Congress of Phonetic Sciences, at which Lacerda publicly presented this new instrument in 1932 (see Figure 1).

Lacerda invented chromography in response to his criticism of the kymograph. Consequently, several of Lacerda's publications on chromography and kymography criticism are available on a digital device. This section also covers other periods in the life and scientific work of the Portuguese phonetician, including his stay at the University of Salvador – in Brazil – from 1956 to 1957. It also covers his time as a researcher and lecturer at the University of Wisconsin–Madison and Queens College of the City University of New York in the 1960s.



Figure 1. Programme of the First International Congress of Phonetic Sciences (Amsterdam, 1932). Source: Arquivo Familiar Paulo de Lacerda.

3 The University of Coimbra Experimental Phonetics Laboratory

This section is dedicated to the Laboratory of Experimental Phonetics in Coimbra, which was founded in 1936 with the primary objective of conducting scientific research into the Portuguese language. During the mid-20th century, various members of the international scientific community considered it to be the most advanced experimental phonetics laboratory in Europe [1]. The laboratory's acquisition of Lacerda's innovative chromograph was fundamental to this perception. For this reason, the decision was made to exhibit the chromograph to the public. It is probably the only surviving copy, as a similar instrument installed at the University of Salvador in 1956–57 was destroyed by fire [12].

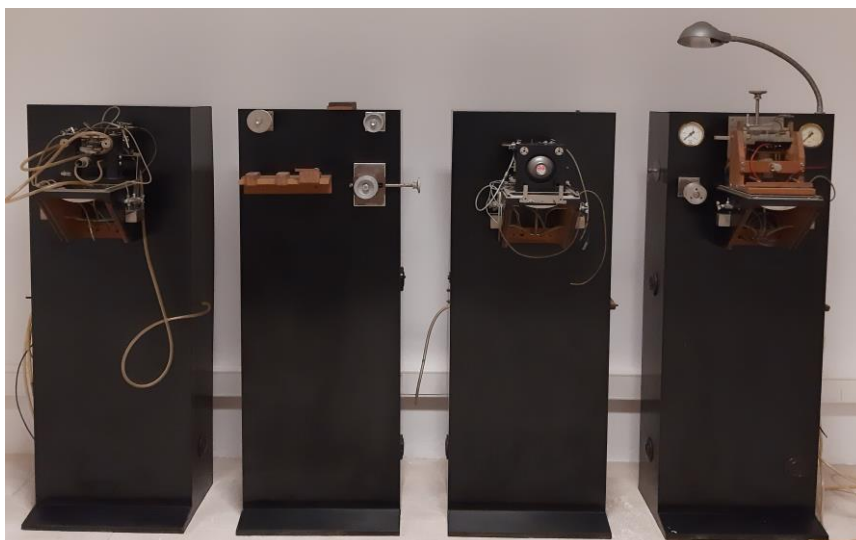


Figure 2. Lacerda's chromograph. Source: Museu da Ciência da Universidade de Coimbra.

Another instrument from the Coimbra Phonetics Laboratory that has been identified in the reserves of the Science Museum of the University of Coimbra and is now on

display is the polychromograph. This is a unique specimen. Alongside the instruments created by Lacerda, there is also an organ that was used in the laboratory's classroom and audition room. Other items include shelves for discs, mouthpieces for the chromograph, a translator of sound configurations into light configurations, and a pitch measuring apparatus. This instrument, used at the Coimbra Experimental Phonetics Laboratory, was built by Constanz Schneider, a laboratory mechanic from Hamburg, who met Armando de Lacerda during his specialist training at the Hamburg Phonetics Laboratory in 1930 and 1931.

As mentioned in [8], it is important to emphasise that the PHONLAB project made it possible to identify, clean, and catalogue all these instruments and objects. The inventory files produced are held by the University of Coimbra Science Museum. “They are also available online on the project website, enabling researchers all over the world to access the information published and providing them with the opportunity to contribute towards its development” [8: 180], [6].

During its existence, the laboratory was housed in two different locations. Initially, it was based in the building of the former Faculty of Letters of Coimbra. In 1951, it moved to new headquarters that were built from scratch as part of a major national public works plan carried out by the Estado Novo (1933-1974). This project was part of the development of the new ‘University City’ of Coimbra – Cidade Universitária de Coimbra, which is now a UNESCO World Heritage Site. Visitors can see the laboratory's first facilities and equipment in the glass photographs of the old Faculty of Letters. A plan for the laboratory shows the facilities of the new faculty and identifies all the rooms that make up the laboratory. Photos of each of the ten rooms that have existed since 1951 are shown, allowing one to see the phonetics laboratory in more detail, including the reading room, the director's office, the phonotheque and archive of graphic documents, the sound archive of Portuguese regional dialects, the microphonic capture chamber for recording sound, the rooms for laboratory work, the room for classifying and analysing sound recordings, the classroom and the audition room.

As the establishment of the Coimbra Phonetics Laboratory involved lengthy negotiations between Lacerda, the University of Coimbra, and the Junta de Educação Nacional (1929–1936) — the first Portuguese institution to plan and fund scientific research, which supported Lacerda's specialisation in Germany from 1930 to 1933 — it was agreed that the exhibition should feature some of the official documentation exchanged between these parties. The aim is to highlight both the interest in introducing experimental phonetics in Portugal and the difficulties that marked its realisation.

4 The international scientific attraction of the Coimbra Phonetics Laboratory

The international prestige of Armando de Lacerda, and the facilities and equipment of the Experimental Phonetics Laboratory explain why, during its existence between 1936 and 1979, the laboratory attracted scientists from universities such as Harvard, Paris, Cambridge, Bonn, Texas, Toulouse, Milan, São Salvador da Bahia, Madrid, Accra, Uppsala, Oslo, Rio de Janeiro, Barcelona, São Paulo, and Edinburgh. These scientists researched and specialised in the innovative phonetic research tools and methods developed by Lacerda. Upon returning to their home institutions, they implemented the techniques, methods, and teachings that they had learned under the supervision of the Portuguese phonetician. This occurred in institutions spanning various countries and continents, ranging from Harvard University, in the USA, to the Federal University of Salvador da Bahia, in Brazil [3].

This section of the exhibition is aimed at the collaborators of the Coimbra Phonetics Laboratory. The development of an interactive map enables visitors to visualise the origins of the phoneticians who came to Coimbra to work at the laboratory. Providing additional information on each of the individuals involved, the map allows visitors to deepen their knowledge of the scientists in question by offering details such as their scientific affiliation, country of origin, length of internship, and the name of the funding organisation.

Papers published by several collaborators of the Coimbra Phonetics Laboratory will be on display, some of which are dedicated to Lacerda. There will be also photographs, letters, and newspaper cuttings that document their presence and activity at the laboratory. This part of the exhibition features names such as Göran Hammarström (1922–2019), Peter Strevens (1922–1989), Francis M. Rogers (1914–1989), Maria Josefa Canellada (1912–1995), and Antoni Badia Margarit (1920–2014), to name a few.



Figure 3. Peter Strevens (left), Armando de Lacerda, and Sue Nogueira at the Coimbra Phonetics Laboratory in the 1950s. Source: Arquivo Familiar Paulo de Lacerda.

The aim of this section is for visitors to realise how young researchers, after a prolonged presence in Coimbra's Experimental Phonetics Laboratory, developed important academic careers at leading universities. The concern to show how research schools were created in other parts of the world from the Coimbra laboratory resulted in the emphasis on the case study of São Salvador da Bahia (Brazil), where Armando de Lacerda and his disciple Nelson Rossi (1927-2014) set up the first experimental phonetics laboratory in South America in 1956-1957. This laboratory was equipped with chromographs, and supported the production of the first linguistic atlas of Brazil [12]. Another noteworthy figure in this section is António Almeida, the Portuguese disciple whom Lacerda hoped would succeed him as director of the laboratory after his retirement in 1972.

For some of these Coimbra Phonetics Laboratory collaborators, this exhibition reinforces the memory that still persists in some institutions. In other cases, however, this exhibition is a place that recovers and preserves the memory of the existence and activity of important linguists and phoneticians who have fallen into oblivion over time.

5 Sound archive of Portuguese regional varieties, 1950s

The exhibition concludes with one of the Coimbra Phonetics Laboratory's key projects: the creation of the sound archive of Portuguese regional dialects. This project,

developed in the early fifties, involved the engagement of a range of individuals who were well-respected in the local area: teachers, doctors, priests, local landowners and archaeologists were some of the figures involved who helped laboratory researchers to select speakers, topics of conversation and recording venues.



Figure 4. In the photograph, from right to left, are Joaquim Roque, a teacher, Berta Lacerda and Armando de Lacerda in a street in Barrancos, in Portugal’s Alentejo region, 1950s. Source: Arquivo Familiar Paulo de Lacerda.

“The appearance of scientists equipped with instruments for observing and recording sounds descending on the far-flung towns and villages of Portugal caused some surprise and anxiety among the local inhabitants, but the presence of Berta, Lacerda’s wife and collaborator, created a climate of trust among women and children, who agreed to her recording their speech. The essential presence of a female collaborator and the aid of local intermediaries enabled “Notes on speakers” and photographs to be added to the sound archive, thus documenting rural Portugal of the 1950s [1: 115]”.

This section highlights notes and photographs of rural workers, accompanied by woodcuts, and newspaper articles from Portuguese and foreign publications reporting on this sound archive.

As this exhibition focuses on the history of experimental phonetics, sound is an essential element. As mentioned in [8], all the reels belonging to the Coimbra Phonetics Laboratory that were found had previously been digitised at the Institute of Phonetics at Stockholm University. Throughout the exhibition, and particularly in the final section, visitors will have the opportunity to listen to recordings of Portuguese phonetics courses for foreigners taught by Armando de Lacerda via sound points. These historical sources have been supplemented by testimonies from former disciples and students of the Portuguese phonetician. The aim is to provide a first-hand account of what it was like to work and learn in this experimental phonetics laboratory.

6 Funding

The IHC is funded by National funds through FCT — Fundação para a Ciência e a Tecnologia, I.P., under the projects UID/04209 and LA/P/0132/2020 (DOI 10.54499/LA/P/0132/2020). Research for this paper was also funded by the FCT under projects 2021.00906.CEECIND and 2022.06811.PTDC (Phonetics Laboratory: Coimbra – Harvard. Rethinking 20th-century scientific centres and peripheries).

7 References

- [1] LOPES, Q.: *The global periphery. Armando de Lacerda and the Coimbra Experimental Phonetics Laboratory (1936-1979)*. Stockholm, The Royal Swedish Academy of Sciences. Center for History of Science, 2023.
- [2] LOPES, Q., A. BRAUN, and M. ASHBY (eds.): *Lacerda 120. Proceedings of the Fifth International Workshop on the History of Speech Communication Research*. Dresden, TUDpress, 2022.
- [3] LOPES, Q., F. LACERDA, and A. SIMÕES: *Armando de Lacerda and the Coimbra Phonetics Laboratory, 1930-1979. Cross-national mobility and exchange in a global context, Centaurus. Journal of the European Society for the History of Science*, 66 (3), pp. 319-350, 2024. doi.org/10.1484/J.CNT.5.149750
- [4] lacerda120.wordpress.com/
- [5] www.kva.se/en/event/armando-de-lacerda-a-pioneer-of-experimental-phonetics-2/
- [6] www.phonlab.uevora.pt/en/home-english/
- [7] Comissão Nacional da UNESCO – unescoportugal.mne.gov.pt/pt/temas/educacao-para-o-seculo-xxi
- [8] LOPES, Q., H. QUARESMA, D. FARIAS, M. CRUZ, and F. LACERDA: *The Phonlab project: Coimbra Phonetics Laboratory – heritage preservation and development*. In: J. BÓNA (ed.): *Proceedings of the Sixth International Workshop on the History of Speech Communication Research*, pp. 177-187, Dresden, TUDpress, 2024.
- [9] MEHNERT, D.: *Historische phonetische Geräte. Katalog der historischen akustisch-phonetischen Sammlung (HAPS) der Technischen Universität Dresden, erster Teil*. Dresden, TUDpress, 2012.
- [10] HOFFMANN, R.: *Historische Objekte der Sprachakustik. Katalog der historischen akustisch-phonetischen Sammlung (HAPS) der Technischen Universität Dresden, zweiter Teil*. Dresden, TUDpress, 2021.
- [11] HOFFMANN, R.: *20 Jahre historische akustisch-phonetische Sammlung (HAPS) der TU Dresden, 1999-2019*. Dresden, TUDpress, 2020.
- [12] MOTA, J., and J. OLIVEIRA: *Armando de Lacerda and the first phonetics laboratory in Brazil*. In: Q. LOPES, A. BRAUN, and M. ASHBY (eds.): *Lacerda 120. Proceedings of the Fifth International Workshop on the History of Speech Communication Research*, pp. 111-117, Dresden, TUDpress, 2022.

IMPACT OF BACKGROUND MUSIC REMOVAL USING DEEP NEURAL MODELS ON FORMANT MEASUREMENTS IN HISTORICAL NEWSREELS CORPORA

Juliusz Cęcelewski¹, Cédric Gendrot¹

¹*Laboratoire de Phonétique et Phonologie (UMR 7018),*

CNRS & U. Sorbonne-Nouvelle, France

[*juliusz.cecelewski@sorbonne-nouvelle.fr*](mailto:juliusz.cecelewski@sorbonne-nouvelle.fr)

Abstract: This study examines the impact of deep neural network models, for background music removal on automated formant extraction accuracy (F1, F2, F3, and F4). The dataset consists of a 1956 Polish news chronicle, with varying audio quality, background noise, and orchestral accompaniment. Two Control Groups included the original and denoised audio, both preserving the background music. The Experimental Groups featured four versions of the same recording, processed with different separation models. Results show higher variability and detection errors in the Control Groups, particularly for the front vowel /i/ and the mid-open nasalized rounded back vowel /ɔ̃/, due to harmonic interference from the background music. In contrast, formant values in the Experimental Groups remained stable across processing methods, with no significant differences among the DNN models used, closely aligning with reference values for Polish vowels.

1 Introduction

1.1 Computational Challenges in Investigating Historical Speech Corpora

The acoustic study of corpora recorded in real-world conditions, both contemporary and historical, provides valuable insights into the phonetics and phonology of speech [8,17]. However, analyses of pre-existing recordings face challenges in generalizability due to uncontrolled parameters and hidden variables. Working with recordings of varying quality often necessitates focusing on a minimal set of robust acoustic parameters, typically excluding most noise-based measures.

Furthermore, in many radio or filmed newsreels from the 1920s to the late 20th century, background music and sound effects overlap with the speaker’s voice—a common practice in, e.g. (among many others) 1930s-1970s French newsreels, American newsreel, (1911-1967) [6], British Movietone News (1929–1979) [19], Soviet Soyuzkinozhurnal (1931–1944) [20], and the Polish Film Chronicle (1944–1994) [3] analyzed here.

Background audio often included orchestration—performed by professional radio orchestras or later generated electronically—as well as other sounds such as singing, applause, nature sounds, urban noise, traffic, industrial machinery, or crowd noise, reflecting the cultural and social context of the broadcast. These diverse audio elements complicate speech analysis, requiring separation from the vocal stem to ensure accurate linguistic and acoustic measurements.

The aim of this study is to compare different processing methods applied to a sample of scripted connected speech to optimize the extraction of formant values (F1, F2, F3, and F4) for Polish oral and nasal vowels /i, ĩ, ε, ɛ̃, a, ɔ, ɔ̃, u/, after separating the audio into a vocal and an instrumental stem to minimize interference from the harmonic structure of the background sounds with the automated extraction of formant values.

We hypothesize that DNN-based background noise removal will reduce formant variability, bringing values closer to those reported in the literature and enhancing the reliability of archival speech for phonetic analysis.

2 Method

2.1 Corpus

A 15-minute archival segment from 1956 was selected from the Polska Kronika Filmowa (Polish Film Chronicle) corpus, a series of newsreels screened in Polish cinemas before the main feature from 1944 to 1994. These newsreels covered current events, the economy, sports, culture, commentary and entertainment, all delivered in a official journalistic style by professional speakers. A summary of the archive’s global acoustic characteristics, as well as a selection of thematic content and background sounds, is provided in Table 1.

Table 1. Selected Thematic Sections and Accompanying Background Sounds from the analyzed recording. Global characteristics of the recording: Title, Total Duration, Mean Amplitude, RMS Energy, Zero-Crossing Rate, and Spectral Rolloff.

Thematic Section	Background Sounds
<i>20th Congress of the Polish Communist Party</i>	Strings, piano, applause
<i>Launch of the Rolling Mill at the Lenin Steelworks</i>	Sounds of electric rolling mill engines
<i>Battling Winter in the Baltic Sea</i>	Strings, brass instruments
<i>The National Library in Warsaw</i>	Flute, bells, strings
<i>The Young Bears at the Warsaw Zoo</i>	Whining of the bear cubs
<i>Fashion Show in Warsaw</i>	Jazz music
<i>Boxing Match: Poland vs. West Germany</i>	Cheers from the crowd around the ring
Title: Polska Kronika Filmowa 10/1956	Total Duration (s): 922
Mean Amplitude: -1.236	RMS Energy: 0.095
Zero-Crossing Rate: 0.038	Spectral Rolloff (Hz): 2466.6

2.2 State-of-the-Art Source Separation Models for Background Music Removal

A variety of commercially available tools are currently available for separating vocal from instrumental stems, such as RX10 [10], SpectraLayers [21], EaseUS Online Vocal Remover [4], PhonicMind [15], MyEdit [1], Media.io Vocal Remover [13], and LALAL.AI Vocal Cleaner [11], among many others. More advanced tools, such as the Demucs project [18] and its derivatives—a state-of-the-art music source separation model—enable even greater levels of separation quality and fidelity.

For this study, we selected Ultimate Vocal Remover v5.6 (UVR) [23], which is an open-source, cross-platform software for separating vocals and instruments from audio files. Compatible with Windows, macOS, and Linux, it supports a wide range of audio formats, including MP3, FLAC, OGG, WAV, and other FFmpeg-supported formats. The latest release (v5.6.0) integrates advanced AI-driven features allowing users to extract vocal and instrumental stems separately.

UVR provides an extensive library of pre-trained models based on two primary architectures: VR and MDX-Net [23]. These models are optimized for a range of tasks, including isolating vocal and instrumental stems, separating lead and backing vocals, decomposing instrumental stems into individual instruments, suppressing background noise, and reducing echo or reverberation.

A distinguishing feature of UVR is its “ensemble” mode, allowing users to combine outputs from different AI models or average multiple outputs to improve the accuracy of audio separation. Although the tool supports Demucs models [18], the recommended architecture is MDX-Net for its superior performance in extracting high-quality vocal stems; however, unlike Demucs, MDX-Net specializes solely in two-stem separation (vocals and other).

The performance of a model in isolating a high-quality vocal stem can be evaluated by its ability to preserve speech components susceptible to confusion, such as vocoids that overlap with the harmonic structure of the music or unvoiced obstruents that resemble aperiodic noise-like signals. Furthermore, optimal isolation should prevent the introduction of artifacts that were not present in the original vocal production.

For this study, two processing groups were established to compare the original unprocessed signal, the same signal denoised with the UVR De-Noise model, and four processing methods utilizing one or more UVR models optimized for high-quality vocal stem separation. The models and their settings are detailed in Table 2.

Table 2. Key characteristics of the separation models used: Process Method, UVR Model code, Window Size, and Output Stems.

Id.	Group	Process Method	Model	Output
1	Control Group	—	—	Original Audio
2	Control Group	VR Architecture	UVR-DeNoise	Denoise Track + Noise Stem
3	Experimental Group	VR Architecture	3-HP-Vocal UVR	Vocal Stem + Instrumental Stem
4	Experimental Group	MDX-Net	MDX23C-InstVoc HQ	Vocal Stem + Instrumental Stem
5	Experimental Group	MDX-Net	Kim-2	Vocal Stem + Instrumental Stem
6	Experimental Group	Ensemble Mode	UVR-MDX-NET 5-HP-UVR	Vocal Stem + Instrumental Stem

2.3 Vocal Stem Separation with UVR

The original audio was processed using five different methods, as detailed in Table 2, to isolate, for the Control Groups, the noise component from the denoised stem, and for the Experimental Groups, the vocal stems from the instrumentals.

Figure 1 illustrates the comparison between two audio processing methods. On the left, we present an excerpt from the original audio processed with the UVR De-Noise model, yielding two stems: the denoised audio with background music and the noise stem alone. On the right, the spectrograms and oscillograms display the vocal and instrumental stems produced by the 3-HP-Vocal-UVR model. This visual comparison highlights the differences between the denoised stem from the control condition and the vocal stem from the experimental group—two versions that show the most pronounced discrepancies in formant extraction accuracy.

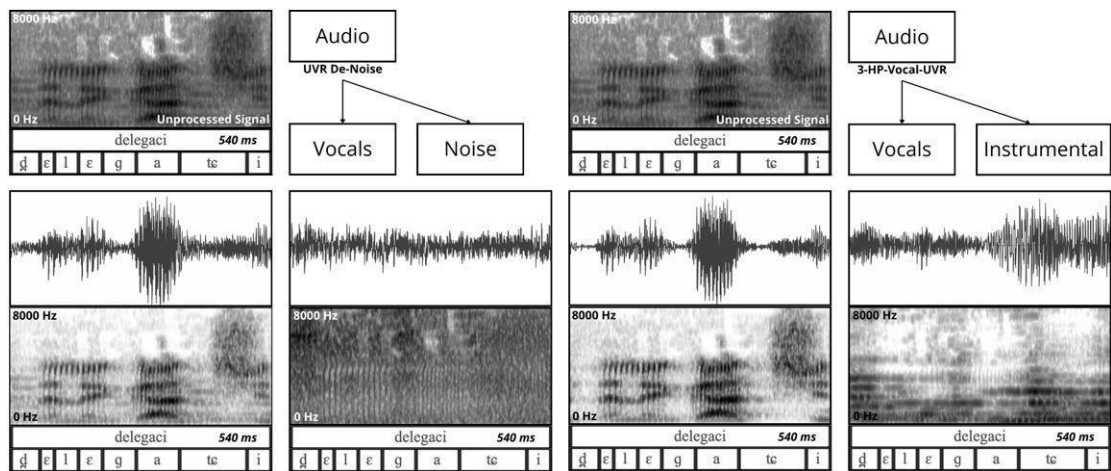


Figure 1. Spectrograms and MFA alignments for the original recording, denoised, and noise stems processed with the UVR De-Noise model (left) and with the 3-HP-Vocal_UVR model (right).

2.4 Data Preprocessing: Orthographic Transcription, Phonemic Segmentation and Alignment

The recording PKF 1956/10 was orthographically transcribed using OpenAI's Whisper based on the vocal stem processed with the MDX23C-Vocal model from the MDX-Net architecture in UVR. An attempt to transcribe the original recording, which included background sounds,

proved ineffective, with an error rate (percentage of incorrect or missing words relative to the total word count) exceeding 40%. In contrast, the error rate for the isolated vocal stem was importantly lower, at 2–3%.

The transcription, including timestamps at the beginning and end of 2-4 sentence portions, was subsequently used to segment into phonemes and align all versions from both the control and experimental groups using the Montreal Forced Aligner (MFA) [22]. The existing Polish pronunciation dictionary was utilized, along with the Polish MFA acoustic model v2_0_0, which was adapted to the dataset under study to improve accuracy.

2.5 The Polish Vowel System and Typical Formant Values

The Polish vowel system comprises six oral phonemes: /i, ɨ, ε, a, ɔ, u/. Historically, it has also been described as including two nasal vowels, [ɛ̃] and [ɔ̃], represented orthographically as <ą> and <ę>, respectively. Polish is the last Slavic language to retain historical nasal vowels [11]. During the 20th century, these phonemes changed from nasal monophthongs to sequences of two (or, according to some studies, three [7]) acoustically distinct segments. The initial segment of these sequences closely resembles the non-nasalized vowels [ε] and [ɔ], respectively. Nevertheless, during the period under investigation (1956), and particularly in the formal and careful speech of professional speakers, nasal vowels can be more accurately described as approximating nasal monophthongs. Contemporary sources from the era [24] provide typical formant values for Polish oral vowels as produced by adult male speakers, which are presented in Table 3 for comparison with the results discussed below.

Table 3. Mean F1 and F2 Values (Hz) for Polish Vowels /i, ɨ, ε, a, ɔ, u/, adapted from [24].

Phoneme	F1 (Hz)	F2 (Hz)
/i/	≈ 250	≈ 2200
/ɨ/	≈ 300	≈ 1900
/ε/	≈ 500	≈ 1900
/a/	≈ 700	≈ 1200
/ɔ/	≈ 500	≈ 1000
/u/	≈ 300	≈ 800

2.6 Extraction of Formant Values

The values of the first four formant peaks (F1, F2, F3, and F4) were extracted at the midpoint of all vocalic segments /i, ɨ, ε, ɛ̃, a, ɔ, ɔ̃, u/. A script was used to automate the extraction process with the Praat software [2]. Lexical stress and unstressed positions were analyzed together. Instead of performing data cleaning after extraction, we applied the formant ceiling optimization procedure described by [5]), ensuring optimal extraction without the need for subsequent outlier removal.

This procedure optimizes the formant ceiling for each vowel. For all tokens, the first four formants were measured 201 times, corresponding to ceilings ranging from 4000 to 6000 Hz in 10 Hz increments. The ‘optimal ceiling’ was determined as the one producing the lowest variation across the measured sets of F1, F2, F3, and F4 values. This variation is calculated on a logarithmic scale as the variance of the $20 \log(F1)$ values, plus the variances of $\log(F2)$, $\log(F3)$, and $\log(F4)$. The procedure results in eight optimal ceilings, one for each vowel.

The extracted values were visualized as vowel charts in the F1 x F2 space (Figure 2) and the F2 x F3 space (Appendix 1), with separate plots for each processed version. Ellipses represent 95% confidence intervals, encompassing data points in different colors referencing their corresponding phonological vowel.

In Appendices 2 and 3, the F1, F2, and, respectively, F3 and F4 values are displayed as boxplots for each vowel across all recording versions.

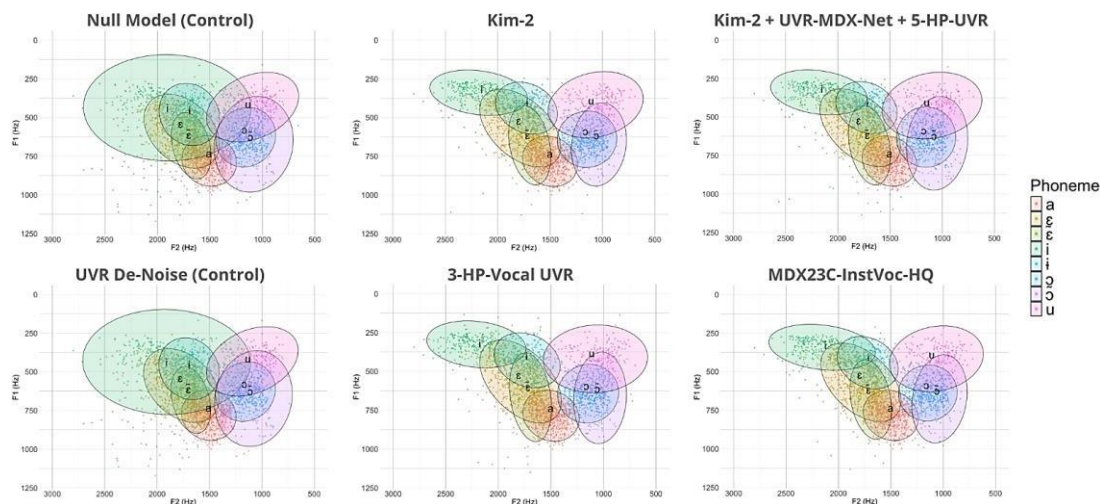


Figure 2. Mean F1 and F2 values (Hz) and 95% occurrence ellipses for the vowel targets /i, i, ε, ε̃, a, ɔ, ɔ̃, u/, extracted from recordings processed with various separation models.

3 Results

3.1 Statistical Modeling with One-Way ANOVA and Tukey Post Hoc Comparisons

The F1, F2, F3, and F4 values measured for the eight analyzed vowels /i, i, ε, ε̃, a, ɔ, ɔ̃, u/ were compared across each version of the audio processed with different models, both within the control/experimental groups and between groups. For each comparison, a one-way ANOVA was conducted, and when at least one comparison between models was found to be significant, a Tukey Post Hoc test was subsequently performed to inspect the individual comparisons. To control for the increased risk of Type I errors due to multiple comparisons, we applied the Bonferroni correction to adjust the significance levels of the post hoc tests. All statistical validations were conducted in the R environment, Version 2024.09.1+394 [16].

Post Hoc tests were conducted for the F1, F2, F3, and F4 of /i, ε, ɔ, u/, the F1 and F3 of /i/, and the F3 of /ɔ̃/, as displayed in Appendices 4, 5, 6, 7, 8, and 9 for the vowels /i, i, ε, ɔ, ɔ̃, u/, respectively. However, since the ANOVA results were not significant for comparisons of F1, F2, F3, and F4 for /a, ε̃/, F2 and F4 for /i/, and F1, F2, and F4 for /ɔ̃/, Post Hoc tests were not conducted for these cases.

3.2 Impact of Processing by DNN Models on Formant Detection

Among the 118 significantly different comparisons, 113 involved the difference between the control and experimental groups, i.e. [Null Model or UVR De-Noise] versus [3-HP-Vocal UVR, MDX23C-InstVoc HQ, Kim-2, or Kim-2 + UVR-MDX-NET + 5-HP-UVR]. Only 5 comparisons showed significant differences between different models within the experimental group. For instance, the F3 and F4 of /i, u/ were significantly higher in the audio processed with the MDX23C-InstVoc HQ model compared to the 3-HP-Vocal UVR model, and the F4 of /u/ was significantly higher in the audio processed with the MDX23C-InstVoc HQ model compared to the combined Kim-2 + UVR-MDX-NET + 5-HP-UVR model. The only notable regularity, which runs counter to the general trends observed in group comparisons, is the significantly greater spread of F4 values for /u/, approximately 200 Hz, in the experimental group compared to the control group.

Figure 3 presents the mean absolute errors calculated for a comparative analysis between the reference F1 and F2 values [24] and our own results. Given the minimal variation within both the experimental and control groups, we only display the comparison between these two groups. Errors were larger in the control group in 8 out of 12 comparisons. In the remaining

four cases, mean error values were nearly identical across both groups, with the vowel /a/ showing the least sensitivity of F1 and F2 values to the presence of background sounds in the analyzed archive.

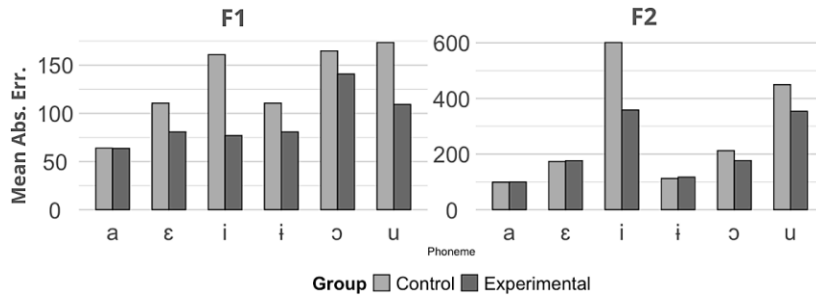


Figure 3. Absolute Mean Error between reference values from [23] and our results for F1 and F2 of /i, i, ε, a, o, u/ in the Control and Experimental Groups.

Figure 4 displays the spectrograms of an excerpt from the original recording (left), where the presence of background music leads to the automatic detection of partially aberrant formants throughout the syllable [pre], including the unvoiced [p]. The formants highlighted in red provide a visual assessment of the extent to which background music distorts spectral measurements and interferes with the formant structure of sonorants, leading to detection errors. By contrast, in the spectrogram of the vocal stem processed with the Kim-2 model (right), the removal of the orchestration restores a more regular formant structure in the sonorants, while [p] regains its characteristic continuous aperiodic signal.

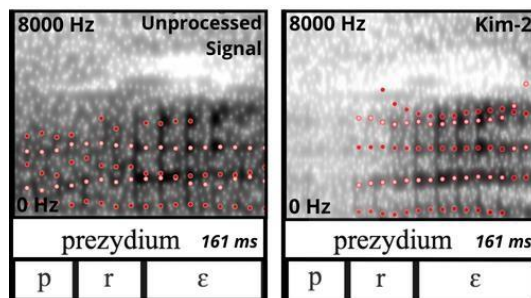


Figure 4. Spectrograms with the first 4 formants highlighted by Praat, and MFA alignments for the original recording (left) and the vocal stem (right) processed with the Kim-2 model.

4 Discussion

The analysis of isolated vocal stems appears to be a reliable and robust solution for corpora originally containing various types of background sounds. All tested models, including a set of models combined using the “ensemble” mode, proved effective, and the extracted formant values do not allow us to determine whether one model outperformed the others. However, we demonstrated that extraction from vocal stems, compared to audio retaining background sounds—whether treated or not with a noise reduction model—is more accurate and constrains the distribution of data points to more clearly defined regions within the vowel space.

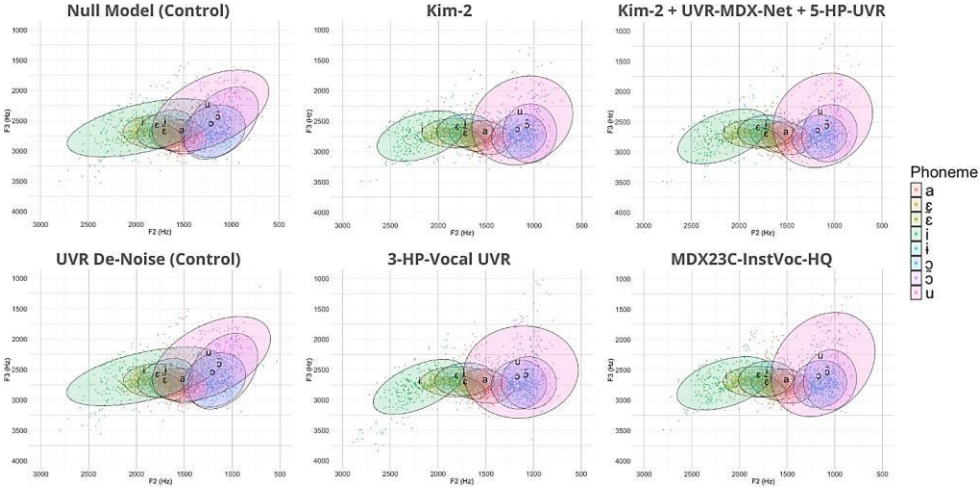
This observation applies to both the oral vowels studied (/i, i, ε, a, o, u/) and the two nasal (or partially nasalized) vowels (/ɛ̃, ɔ̃/). While the degree of nasalization in Polish vowels remains a research question beyond the scope of this study [7], it is worth noting that initial predictions suggested that extracting formant values from nasal or nasalized vowels would result in a higher number of erroneous or implausible values compared to oral vowels. This was not observed in the analyzed corpus.

In summary, the results support recommending the removal of background sounds when present in audio corpus data, including archival recordings. We are confident that the method described here will enable the use of valuable linguistic data that has traditionally been excluded due to the superposition of non-speech sounds.

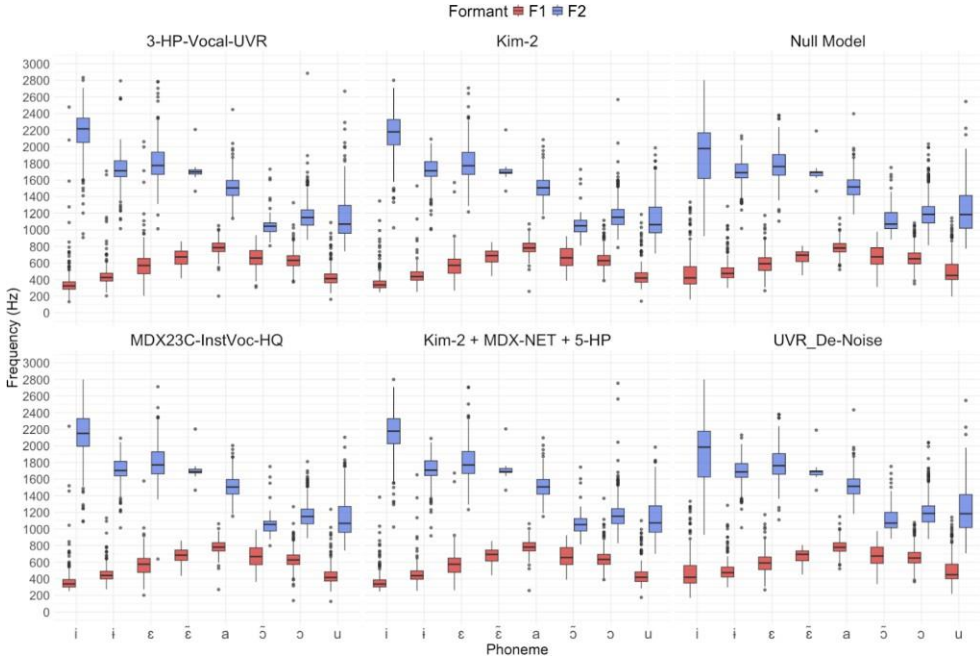
5 References

- [1] ADOBE. (n.d.). *MyEdit AI Vocal Remover*. <https://www.adobe.com/express/feature/audio/remove-vocals>
- [2] BOERSMA, P. and WEENINK, D. (2016). PRAAT: Doing phonetics by computer [Computer program]. Version 6.4.25. Retrieved 10 Dec 2024, from www.fon.hum.uva.nl/praat.
- [3] CIEŚLIŃSKI, M. (2006). *Piękniej niż w życiu. Polska Kronika Filmowa, 1944-1994*, TRIO.
- [4] EASEUS. (n.d.). *EaseUS Online Vocal Remover*. <https://www.easeus.com/online-vocal-remover.html>
- [5] ESCUDERO, P., BOERSMA, P., RAUBER, A. and BION, R. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America*. 126. 1379–1393.
- [6] FIELDING, R. (1972). *The American newsreel, 1911-1967*. Norman: University of Oklahoma Press.
- [7] GUBRYNOWICZ, R. (2016). Acoustic analysis of Polish nasalised and nasal vowels. *Poradnik Jezykowy*. 7–18.
- [8] HARRINGTON, J., Paethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: an acoustic analysis of the Queen’s Christmas broadcasts. *Journal of the International Phonetic Association*, 30(1/2), 63–78.
- [9] HURET, M. (1984). *Ciné actualités : histoire de la presse filmée, 1895-1980*, Paris, H. Veyrier, 192 p.
- [10] IZOTOPE. (2023). *RX10 Audio Editor* (Version 10). iZotope Inc. <https://www.izotope.com/en/products/rx.html>
- [11] KONECZNA, H. (1965). *Charakterystyka fonetyczna języka polskiego na tle innych języków słowiańskich*. Warsaw, PWN.
- [12] LALAL.AI. (n.d.). *LALAL.AI Vocal Cleaner*. <https://www.lalal.ai>
- [13] MEDIA.IO. (n.d.). *Media.io Vocal Remover*. <https://www.media.io/vocal-remover.html>
- [14] OPENAI. (2022). *Whisper* [Speech recognition model]. <https://github.com/openai/whisper>
- [15] PHONICMIND. (n.d.). *PhonicMind AI Vocal Remover*. <https://phonicmind.com>.
- [16] R CORE TEAM (2021) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org>.
- [17] RATHCKE, T., STUART-SMITH, J., TORSNEY, B. and HARRINGTON, J. (2017). The beauty in a beast: Minimising the effects of diverse recording quality on vowel formant measurements in sociophonetic real-time studies. In: *Speech Communication*, Vol. 86: pp. 24-41.
- [18] ROUARD, S., MASSA, F., & DEFOSSEZ, A. (2023). Hybrid transformers for music source separation. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5).
- [19] SMITHER, R., KLAUE, W. (eds) (1996). *Newsreels in film archives : a survey based on the FIAF Newsreel Symposium*. Wiltshire, England: Flicks Books ; Madison : Fairleigh Dick- inson University Press.
- [20] SOVIET SOYUZKINOZHURNAL. (1931–1944). *Soviet Soyuzkinozhurnal* [Newsreels]. Soyuzkino.
- [21] STEINBERG. (2023). *SpectraLayers Pro 10* (Version 10). Steinberg Media Technologies GmbH. <https://www.steinberg.net/spectrallayers/>.
- [22] MCAULIFFE, M., SOCOLOF, M., MIHUC, S., WAGNER, M. and SONDEREGGER, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th Conference of the International Speech Communication Association*, pp. 498–502.
- [23] UVR (Ultimate Vocal Remover). (2023). *Ultimate Vocal Remover v5.6* [Software]. Available at: <https://github.com/Anjok07/ultimatevocalremovergui>.
- [24] WIERZCHOWSKA, B. (1980). *Fonetyka i fonologia języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich.

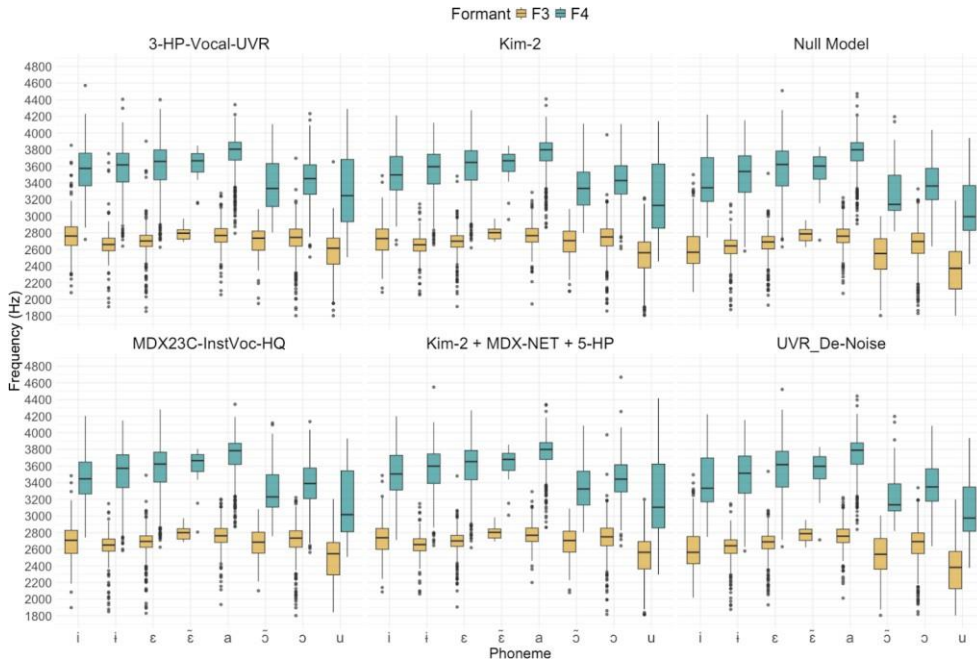
6 Appendices



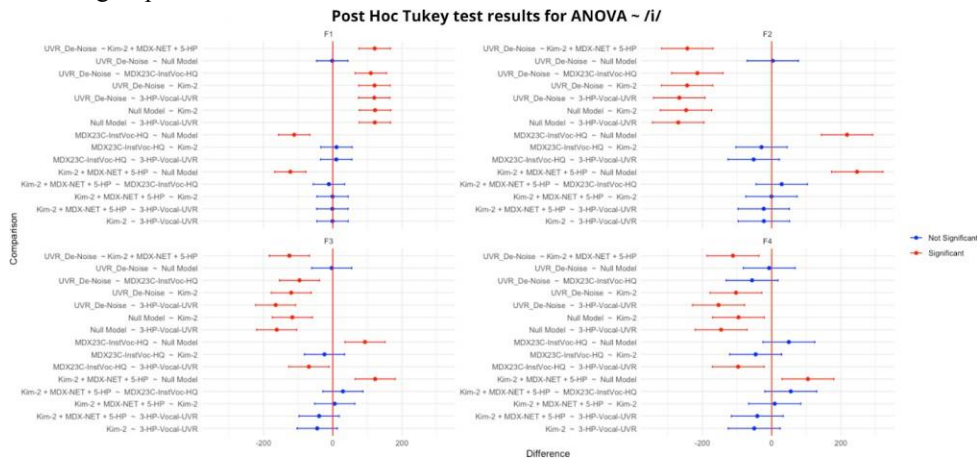
Appendix 1. Mean F3 and F4 values (Hz) and 95% occurrence ellipses for the vowel targets /i, i, ε, ɛ, a, ɔ, ɔ̃, u/, extracted from recordings processed with various separation models.



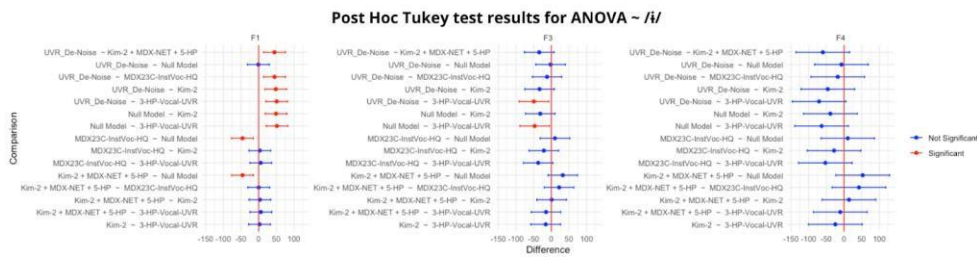
Appendix 2. Box plots of F1 and F2 for /i, i, ε, ɛ, a, ɔ, ɔ̃, u/ by neural model used for the control and experimental groups.



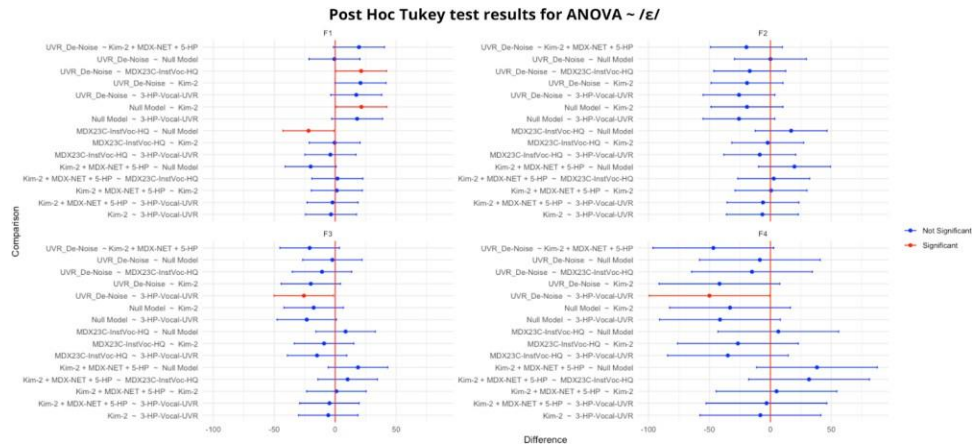
Appendix 3. Box plots of F3 and F4 for /i, i, ε, ε̃, a, ɔ, ɔ̃, u/ by neural model used for the control and experimental groups.



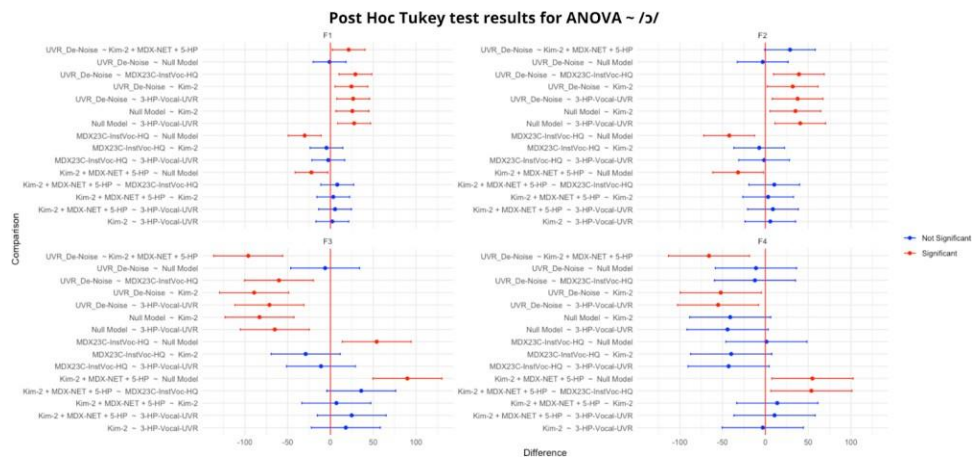
Appendix 4. Post-hoc Tukey test results for ANOVA comparisons by formant (F1, F2, F3, F4) for the vowel /i/.



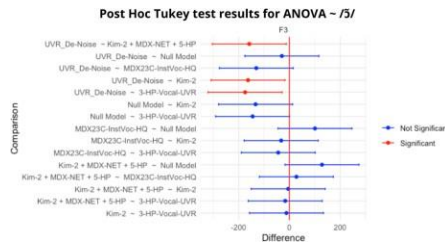
Appendix 5. Post-hoc Tukey test results for ANOVA comparisons by formant (F1, F3, F4) for the vowel /i/.



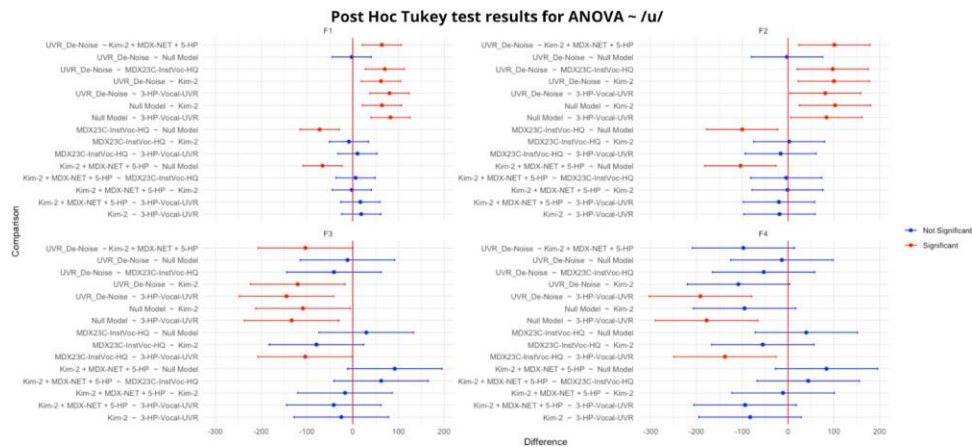
Appendix 6. Post-hoc Tukey test results for ANOVA comparisons by formant (F1, F2, F3, F4) for the vowel /ε/.



Appendix 7. Post-hoc Tukey test results for ANOVA comparisons by formant (F1, F2, F3, F4) for the vowel /ɔ/.



Appendix 8. Post-hoc Tukey test results for ANOVA comparisons for F3 of the vowel /ɜ/.



Appendix 9. Post-hoc Tukey test results for ANOVA comparisons by formant (F1, F2, F3, F4) for the vowel /u/.

SPEECH SITUATIONS IN OLD SOUND NEWSREELS

Gocsál, Ákos^{1,2}

¹University of Pécs, Faculty of Music and Visual Arts, Pécs, Hungary

²HUN-REN Hungarian Research Centre for Linguistics, Budapest, Hungary
gocsal.akos@pte.hu

Abstract: Analyses of the spoken material of old sound newsreels offer great potential for insight into the way people spoke many decades ago. However, from a researcher's point of view, newsreels represent a highly diverse and unstructured body of speech recordings, making it difficult to identify the best recordings for individual studies. Previous research highlighted the importance of using databases for this purpose. A crucial aspect of database development is the establishment of a metadata structure, which, providing keywords for searching recordings, facilitates research by the identification of the recordings for a given research project. The present study examines the possibility of using some of the factors described by Dell Hymes in his model of speech communication for metadata structure development. Qualitative analysis was performed on the spoken material of the news stories produced in 1931 and 1932 in Hungary, with a focus on speech situations, speech events and participants. The research revealed a variety of categories belonging to these factors. It is concluded that the factors Hymes described may provide a good basis for metadata structure development; however, these factors seem to be too general, and addition of subcategories is needed for a more accurate description of speech events.

1 Introduction

Sound newsreels are authentic recordings of how people spoke and used language many decades ago. While there are several newsreel archives available online, such as British Movietone [2], British Pathé [3], or Hungarian newsreel archive [9], only few studies have addressed research questions related to language use in newsreels. A previous study by the present author revealed that some of those studies were conducted by historians or media researchers [10]. The reason why speech is important for those researchers is that it has a significant role in constructing meaning. For example, in a Spanish newsreel, the role of voice-overs was contextualization, but there were parts in a newsreel where testimonies of individual characters were dominant, making the message of the news story more convincing [20]. Other media researcher illustrated voice-overs in quotations, like ... *the voice-over remarks "Winston Churchill's face looks reproachfully..."* [4] or the voiceover states, "*so the door plates featuring names of African countries..."* [24], or speech acts, like *the voice-over praises* [4], *notes that, wittingly asks* [28]. However, it is not only the verbal message that is important in constructing meaning. The tone of voice was as important as the message itself. As Schwarz says, "the sensual quality of the voice carried social and political qualities" [22]. The use of different vocal qualities contributed to a grand narrative, which the newsreels provided, for example, an optimistic tone during the period of reconstruction of Germany after WW2 [13]. Another study remarks that in Spanish newsreels narrators often quoted or paraphrased other characters' texts, and an ironic tone gave an ironic twist to the news story [20]. Voice was perceived as authoritative in the early British Movietone newsreels [5]. In the Spanish newsreels produced between 1977 and 1980, an ironic tone was used when the narrator was speaking about Francoism and its legacy, thus distancing himself from and discrediting Franco's era. In other cases, the tone of voice was used to gain trust [20]. In a Swedish newsreel of

1945, an authoritative and formal tone was used to reassure the audience that the rescue missions presented are “trustworthy” [1: 175]. Narrative mistakes and mispronunciations were also reported [5]. In some cases, voice-overs mention speakers like Franco or Count Folke Bernadotte [1], [10]. The language of the voice-overs is also mentioned in media studies, for example, French [4] or Swedish [1: 178].

The papers mentioned so far only discuss some observations of language use without the application of research methods used by linguists. Other researchers, however, collect linguistic data in a systematic way and analyze them using well established empirical methodologies. The few examples include, in addition to those cited in [11], an analysis of formant changes observed in Danish news reading. Realizations of /æ/ were found to move to a more frontal position, while /ʌ/ was fronted in the 1970s, then backed in the 1990s, compared to the 1950s [27]. Another study suggested a pragmatic approach for newsreel research. Since the producers of the newsreels clearly had intentions, application of the speech act theory is suggested in [18] for research of newsreel commentaries and original sounds like talks or interviews. The purpose of the present paper is similar: Developing a theoretical framework that can provide relevant aspects for researching speech in newsreels.

2 Databases in newsreel research

Newsreel archives contain great amounts of spoken material, offering a wide range of utterances for a variety of research questions. However, due to this diverse nature of the newsreels, identification of the best utterances for a research project may be challenging. Databases, built specifically for this purpose, may facilitate finding the appropriate spoken material. For the development of such databases, Drude et al. offer useful guidelines [7]. Their work describes three types of data that can be distinguished regarding databases of spoken utterances. Primary data are the recordings themselves that contain the spoken material. Secondary data are the annotations, i.e. the explicit representations of the relevant properties that may be relevant for the researcher. These include, but are not limited to, transcriptions and labels of speech phenomena, such as phonetic, morphologic, syntactic, semantic, or pragmatic properties, pauses, or even false starts, corrections or other disfluencies. Annotations in multimodal corpora may also include non-verbal and paralinguistic channels. Secondary data are therefore time-dependent. Finally, effective use of the corpus requires a well-designed set of metadata. Metadata are time-independent descriptors of a whole recording. In other words, metadata translate the news story into a language of keywords [16]. A database of Spanish TV shows is a good example. In this database metadata were available to describe the scene of the news story, such as environment (countryside or city), place (indoor or outdoor), screen (multiple screens shown or not), season (warm or cold weather), time (daylight or nightlight scenes) [19]. Another example, a corpus of Norwegian parliamentary speeches, includes metadata of speakers, i.e. name, gender, date of birth, place of birth, region of birth, electoral district, and dialect [26].

Although these examples pertain to present-day recordings, these approaches can also be used for researching historic newsreels. While most of these metadata mentioned are useful and crucial, they seem to be too general and may not, or may not systematically, address the contextual richness of newsreels for linguistic research. In a previous paper, a possible metadata structure was suggested to facilitate research into old Hungarian newsreels [11]. The metadata included descriptors of situations, stories, speaking styles, and Booleans indicating the presence of different sentence types, dialects, disfluencies, pathologic and emotional speech. Other Booleans indicated the presence of speech, singing, noises, distortions, several speakers etc. Although this metadata structure can be effectively used for filtering recordings, it was not established using a theoretical framework, therefore it may lack crucial features. It is therefore the purpose of this paper to find a comprehensive and coherent system that may

be applied to the description of the newsreels and based on that system, metadata can be derived to label identify speech events and other spoken material more effectively for further research.

The ethnographic approach to speech situations, proposed by Dell Hymes, offers a well-established theoretical background for this purpose. Since the aim here is to support the development of a set of metadata, the focus is on the circumstances of the interactions presented in the newsreels, rather than examining grammatical forms or words. Hymes's idea was to support structural analysis of speech events using a framework describing communicative behaviour [14: 22]. He also defined crucial concepts, such as speech situation, speech event, speech act and speech style. According to this definition, speech situations are general contexts in which speech events occur. Speech situations are not governed by rules of speaking. Hymes' examples include ceremonies, fights, hunts, lovemaking, etc. Speech events are activities directly governed by the rules of speaking. Hymes mentions that a party can be a speech situation, and a conversation in a party can be a speech event. Finally, speech acts are the minimal terms in a speech event, having a status of a command, presented in a variety of forms [15: 56–57]. Hymes also developed a detailed set of factors (the *SPEAKING* model) that can be used for the description of speech events. These are the following:

S: setting and scene – time, place, physical circumstances of a speech act, the cultural definition of an occasion

P: participants – speaker, addressor, hearer, audience, addressee

E: ends – including purposes, outcomes, goals

A: act sequences – message forms and message contents

K: key – tone, manner, spirit

I: instrumentalities – channels and forms of speech, linguistic code

N: norms – specific behaviors related to speaking and interpretation

G: genres – categories representing distinct formal categories traditionally recognized.

Several authors have applied methodologies based on Hymes' model for analyses of speech events in a variety of audio-visual materials. A study thoroughly investigated one of Joe Biden's television speeches, providing aspects related to the *SPEAKING* model [8]. Other researchers analyzed speech events from the tv series *Friends* and demonstrated how they were affected by the factors of the *SPEAKING* model [30]. Other authors used the same model for the analysis of 10 scenes from the movie *Fist Fight* [25]. *SPEAKING* is not the only model that can be used for the description of speech events. Searle's taxonomy of speech acts, such as assertives, directives, commissives, expressives and declarations, is also used for the analysis of utterances [23]. Examples include the analysis of a Walt Disney movie [29], and the feature films *Hive* [21] and *Turning Red* [6]. In the latter example, qualitative methodology was applied to reveal the communicative behavior of the characters.

The cited papers, however, applied these models for an in-depth understanding of certain speech events, and not for supporting the development of a large database. The present paper attempts to apply some of the factors Hymes described to develop categories for a database of metadata.

3 Development of categories based on Hymes's factors

3.1 Materials and methods

In this study, Hungarian sound newsreels, produced between September 1931 and December 1932, are analyzed. Here, "newsreel" is defined as a complete news program, created on a weekly basis, including several different news stories. A "news story" is an individual report

on a particular event. A complete newsreel consisted of 1-10 news stories, most frequently 7 or 8, making up a news program of 10-12 minutes. For the identification of the news stories, the *Newsreel issue no./news story within the newsreel* format is used. For example, 402/2 refers to news story No. 2 in newsreel Issue 402. Issue numbering for sound newsreels started at 398 in September 1931, newsreels with issue numbers below 398 being silent.

From 1931 and 1932, 60 sound newsreels are available, which include 441 news stories. Speech was found in 254 news stories. Languages and number of occurrences in the news stories are as follows: Hungarian 148, German 53, English: 31, French 19, Italian 1, Turkish 1, Bulgarian 1, other 1. Most of the news stories in which foreign languages were spoken were taken from foreign news agencies. In this study, all news stories that contained speech were categorized, regardless of language. News stories with unintelligible utterances, cocktail party situations, etc., were excluded.

A content analysis has been performed to identify wider categories regarding the following:

- a) speech situations
- b) speech events
- c) participants

Each news story in which intelligible speech was found was thoroughly observed and assigned to a category. In this process, a data-driven, inductive approach was applied, that is, categories were not defined previously, and they were continuously revised and restructured as needed when another news story was analyzed [17]. When the news stories were in other languages, the situations visually presented, as well as the title of the news stories, provided enough information to assess the categories.

3.2 Results

3.2.1 Speech situations

The 254 news stories presented a wide range of speech situations. Typical examples include statue inauguration ceremonies [31], the presentation of a new factory [54] or sports events [49]. During the process of analysis, typical situations were identified, and comprehensive categories were formed. Table 1 includes the categories and situation types revealed by the content analysis, with the number of occurrences of news stories in the given category.

Table 1. Speech situation categories and the number of instances. The number of foreign language news in brackets.

No.	Category and situations	No. of instances
1	<u>sports and human achievements</u> : races, trainings, matches, records	60
2	<u>ceremonies</u> : inauguration of statues, funerals, memorial events	33
3	<u>entertainment, leisure time</u> : people dancing, entertaining performances and exhibitions, leisure time activities	26
4	<u>industry, trade, agriculture</u> : factories, exhibition of products, trade, transportation of goods	18
5	<u>foreign politics</u> : foreign politicians' visits, assemblies, rallies, events abroad	18
6	<u>people at work</u> : professionals performing a variety of jobs	13
7	<u>military</u> : parades, training, meetings, military technologies,	12

	soldiers performing tasks	
8	<u>traditions</u> : carnivals, folk traditions	11
9	<u>domestic politics</u> : inauguration of politicians, rallies	8
10	<u>society, social policies</u> : activities related to supporting those in need	8
11	<u>accidents, catastrophe</u> : natural disasters, accidents of vehicles	7
12	<u>culture</u> : artistic performances, artists, actors	7
13	<u>science</u> : presentation of discoveries and experiments	6
14	<u>education</u> : student activities	5
15	<u>technical inventions</u> : presentation of new vehicles and other instruments	5
16	<u>fashion</u> : fashion shows, Miss Universe	5
17	<u>lifestyle</u> : wedding ceremonies, people in everyday situations	5
18	<u>nature, zoo</u> : presentation of images of animals or plants	4
19	<u>new technologies</u> : introduction of new technologies	3

3.2.2 Speech events

In the 254 news stories, 354 speech events were identified. The reason why the number of speech events is considerably larger than that of the news stories is that a high number of news stories included several speech events. In 416/8, a narrator explains that the location is the Hellabrunn Zoo, then the audience hears the spontaneous talk of visitors [39]. Similarly, in 417/5, the narrator informs the audience that in a German town, Mittenwald, citizens make violins in the winter, then the conversation between a violin-making man and a child is displayed [40]. In situations like these, two or more speech events were coded. This observation and practice have led to the following guideline for coding: If, within the same speech situation, there is a change in at least one of the SPEAKING factors, separate speech events should be coded. This may occur even in cases when the same speaker talks to the same audience, but there is a change in genre. For example, in 449/1, a storyteller tells a tale to children. In the first few sentences, the storyteller addresses the audience and tells them that they will be not only listeners but participants of the tale he is going to tell, and who the main characters are, and only then starts telling the tale [53]. Here, the first part of the recording can be coded as “announcement”, and the second part is “storytelling”, or “artistic performance” if larger categories are used. Or, in 418/1, the inauguration ceremony of a statue is shown. The speaker gives a speech about the well-known person, whose statue is inaugurated, then two scouts place a large wreath on the plinth of the statue. In this moment, the speaker tells the scouts how to put the wreath on the plinth [41]. Again, all utterances are from the same speaker, and there is a change in genre (“public speech” and “commands”), but in this case, there is a difference in the participants as well. First, he talks to the audience of the speech, and second, to the scouts, so the addressee of the utterances is different. These examples support the idea that these speech events should be coded separately. It must be noted that in both of these news stories, other factors of the SPEAKING model may also differ in the separate speech events, especially the Ends.

The results of the content analysis of speech events are presented in Table 2. Speech events are general categories and an individual recording where a speech event is identified may comprise several speech acts or only one, but speech acts are not discussed here.

One hundred voice-overs were found in the news stories. Voice-overs or narrations are in general declarations of facts in a variety of contexts like sports events, e.g. the Henley Royal Regatta [51] or nature, e.g. a bird's nest found on a sculpture in Berlin [48]. In one news story from France, narration included a dialogue, in which details of new clothes are discussed in the form of questions and answers between a male and a female speaker [33].

Public speeches represent another large category of speech events, with 65 instances. In public speeches speakers address and talk to a smaller or larger audience. These speech events are monologues and typically occur in ceremonies like inauguration of monuments [55]. Some of the public speeches were read from notes (a speech delivered in memory of Washington [43]) Radio speeches such as [58] or talking to the camera [50] represent a special type within this category.

Also, a relatively high number of conversations (61) was found in the news stories. These include interviews (with a mover [45]) or everyday situations (man wants to buy a Christmas tree [34], policeman shows the way in French and English, then the governor talks to a lady who sells her folk-art products [48]). Naturally occurring, spontaneous situations were also recorded. In 454/2, Harold Lloyd says he would prefer to give an interview at the hotel, but he is informed that the camera is already rolling so he is asked to talk. Also, in the same news story, the news reporter tells the camera operator to start recording when he waves his hand [57]. In a few cases, highly formal conversations were found (military or police officer reports to a commander [32]).

Another large group of speech events (57) includes a variety of commands in which a participant of the situation orders another participant to perform a certain action or activity. Typical examples include sports events when a trainer or a referee instructs participants. Examples include training for skiing in Germany [36], military events, parades, trainings (horse riding of young soldiers in [38]), people at work (disembarking of a ship in [60]) or work with animals (work with bears in Vienna [43]). In some cases, speech events belonging to this category included only one speech act, which was the actual command. In other cases, several speech acts were included, like a short explanation of the command, but the purpose of the whole event was still the execution of certain commands or instructions.

Thirty-three instances of events were identified as announcements, where a speaker explains the context of an event or provides important information. For example, a swimmer explains that they can prepare for the Olympic Games in a swimming pool in the winter [44]. Announcements are similar to public speeches in a way that the announcer is physically present at the event, but their function is similar to that of the voice-overs.

Table 2. Speech event categories and the number of instances.

No.	Category and events	No. of instances
1	voice-overs: narrations in studio or in situ, in dialogue	100
2	public speeches	65
3	conversations	61
4	commands, instructions	57
5	announcements	33
6	speech in crowd	28
7	artistic performance	5

28 recordings included speech events, which can be described as “speech in crowd”. One type of this kind is when members of a group spontaneously utter the same phrase, like *Éljen!* (*hooray or long live!*) expressing their support to a speaker who opens an exhibition [46]. The formal form of speech-in-crowd is oath taking [52]. The third type belonging to this category is when individuals of a group produce utterances separately, like a football supporter shouting *Ébredjete!* (*Wake up!*) [56]. The fourth type of this event is when people in an elevated voice offer their goods for sale like in [59].

There were five instances of artistic performances, mentioning only an actor playing [35], poem recitation [37], and storytelling [53].

3.2.3 Participants

For a better and more accurate description of speech situations, it is crucial to identify and describe the speakers and listeners of the speech events. Table 3 summarizes the findings.

The 100 voice-overs represent the simplest case, since the narrators are the speakers, and the movie audience is in the listener’s role. The 65 cases, where public speeches were found, more attention is required. In all these cases, speakers talked to local audiences. In 42 speech events, speakers were identified (e.g. Géza Ziegler in [31]), therefore, demographic and other important data may be available for those speakers. Thus, when designing a database of metadata, this possibility should be considered.

Participants in conversations are in symmetric roles as all of them can be speakers and listeners too. A great variety of participants were found in conversations. Examples include actor and journalist [57], customer and seller [34] or old man and child making violins [40].

In the commands category, speakers are usually trainers, referees, team leaders, military officers. Listeners and those who execute the commands are athletes, workers or other military persons. In the announcements category, a wide range of speakers was identified, while the listeners are always people present locally.

In the “speech in crowd” category, speakers are members of a group, and in the artistic performance group, speakers were actors and presented their talk either locally, or to the movie audience.

Table 3. Speakers and listeners of the speech events and the number of instances.

No.	Speaker	Listener	No. of instances
1	<u>voice-overs</u> : narrator	movie audience	100
2	<u>speeches</u> : public speakers	local audience	65
3	<u>conversations</u> : actor-journalist, customer-seller etc.		61
4	<u>commands</u> : trainer, referee, team leader etc.	sportsmen, workers etc.	57
5	<u>announcements</u> : announcers, trainers, speakers	local audience	33
6	<u>speech in crowd</u> : members of a group	local audience	28
7	<u>artistic performances</u> : actors	movie or local audience	5

4 Conclusions

As a general conclusion, it can be stated that the categories created based on Hymes's concept may provide a good basis for the development of metadata, thus, keywords derived from the categories presented here may help researchers select or separate recordings based on linguistically relevant factors. It should be noted though that the factors originally defined by Hymes seem to be too broad. For a more accurate identification of recordings in a database, subcategories need to be introduced, facilitating the development of a more detailed keyword system.

In addition to those categories, quantitative indicators, i.e. the number of recordings belonging to the categories were also established. This outcome of the study provides a general description about the corpus of the recordings and may be used for further research on how different aspects of the genre of newsreels, including language use, changed or evolved over time.

There are several limitations of this study. First, only a few factors described by Hymes were tested. Further studies should address the remaining factors of the SPEAKING model, as well as those related to speech acts, so that a more comprehensive system of categories can be developed. Second, the categories developed here were not converted to keywords and tested in an actual database. This is also a task for future work. Nevertheless, it is believed that a metadata system, based on theoretical considerations as demonstrated here, may contribute to the development of a more effective database of metadata and thus help us better understand language use in old newsreels.

5 References

- [1] BERGSTRÖM, Å. and M. JÖNSSON: *Screening War and Peace: Newsreel Pragmatism in Neutral Sweden, September 1939 and May 1945*. In: C. CHAMBERS, M. JÖNSSON and R. V. WINKEL (eds.): *Researching Newsreels* (Global Cinema), pp. 157–182, Springer International Publishing, Cham, 2018. https://doi.org/10.1007/978-3-319-91920-1_9
- [2] British Movietone online archive: <https://www.youtube.com/@BritishMovietone/videos>
- [3] British Pathé online archive: www.britishpathe.com
- [4] CENTENO M. and M.PABLO.: *Reediting the war in Asia: Japanese newsreels in Spain (1931-1945)*, *L'Atalante. Revista de Estudios Cinematográficos* 29. pp. 101–119. 2020.
- [5] CROSSON, S. and D. MCANALLEN: 'CROKE PARK GOES PLUMB CRAZY': *Gaelic Games in Pathé Newsreels, 1920–1939*, *Media History*, 17(2). pp. 159–174, 2011. <https://doi.org/10.1080/13688804.2011.554728>.
- [6] DIFFANI, H. and A. KHOLIS.: *An Analysis of Speech Act in the Movie „Turning Red”*, *Journal of Linguistics, Culture and Communication*, 1(2), pp.100–115, 2023. <https://doi.org/10.61320/jolcc.v1i2.100-115>
- [7] DRUDE, S., P. TRILSBEEK, H. SLOETJES, and D. BROEDER: *Best practices in the creation, archiving and dissemination of speech corpora at the language archive*. In: S. RUHI, M. HAUGH, T. SCHMIDT and K. WÖRNER (eds.): *Best Practices for Spoken Corpora in Linguistic Research*, pp. 183–207, Cambridge Scholars Publishing, Newcastle upon Tyne, 2014. <https://www.mpi.nl/publications/item2051339/best-practices-creation-archiving-and-dissemination-speech-corpora>
- [8] ELZAGHAL, F. T.: *Fields of Text Studies and Discourse Analysis: A Study in Light of Applying Dell Hymes Speaking Model to Joe Biden's Speech on the Middle East (2021)*, *International Journal of Humanities and Language Research*, 4(1), pp. 1–39, 2021. <https://doi.org/10.21608/ijhrlr.2021.225319>
- [9] Filmhíradók online Hungarian newsreel archive: <https://filmhíradokonline.hu/>
- [10] GALLO, G. D.: *Politics and Political Alterity in the Spanish NO-DOs of 1968*. *On_Culture: The Open Journal for the Study of Culture* 4. 2017. <https://jilupub.ub.uni-giessen.de/server/api/core/bitstreams/4f21bf2b-eea0-4f64-a6d5-077b6f4fba8f/content>

- [11] GOCSÁL, Á.: *Voices of the Past: Research into Old Hungarian Newsreels*. In: J. BÓNA (ed.): *HSCR 2024 Proceedings of the Sixth International Workshop on the History of Speech Communication Research*. Budapest, Hungary 20-21 September 2024. Studentexte Zur Sprachkommunikation, 109. pp. 136–45. TUDPress, Dresden. 2024. <https://doi.org/10.21437/HSCR.2024-11>
- [12] HEUVEL, H. VAN DEN, E. KOMEN, and N. OOSTDIJK: *Metadata Collection Records for Language Resources*. In: N. CALZOLARI, K. CHOUKRI and C. CIERI (eds.): *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC2018)*, May 7-12, Miyazaki, Japan, pp. 1282–1288, 2018. <https://hdl.handle.net/2066/192404>.
- [13] HICKETHIER, K.: *The Creation of Cultural Identity through Weekly Newsreels in Germany in the 1950s: As Illustrated by the NEUE DEUTSCHE WOCHENSCHAU and the UFA-WOCHENSCHAU (With a Side Glance at the DEFA Weekly Newsreel DER AUGENZEUGE)*. In: K. IMESCH, S. SCHADE, and S. SIEBER (eds.), *MedienAnalysen*, vol. 17, 1st edn., pp. 39–54, transcript Verlag, Bielefeld, 2016. <https://doi.org/10.14361/9783839429754-003>.
- [14] HYMES, D.: *The Ethnography of Speaking*. In: T. GLADWIN and W. STURTEVANT (eds.): *Anthropology and Human Behavior*. pp. 39–53, Washington: Anthropological Society of Washington, 1962.
- [15] HYMES, D.: *Models of the Interaction of Language and Social Life*. In: J.J. GUMPERZ and D. HYMES (eds.): *Direction in Sociolinguistics: The Ethnography of Communication*, pp. 35–71, Holt, Rinehart and Winston, New York, 1972.
- [16] IBRUS, I. and M. OJAMAA: *Newsreels versus Newspapers versus Metadata: A Comparative Study of Metadata Modelling the 1930s in Estonia*, *VIEW Journal of European Television History and Culture* 7(14). pp. 123–137, 2018. <https://doi.org/10.18146/2213-0969.2018.jethc157>.
- [17] KUCKARTZ, U.: *Qualitative Text Analysis: A Systematic Approach*. In: G. KAISER and N. PRESMEG (eds.): *Compendium for Early Career Researchers in Mathematics Education*, pp. 181–97, Springer International Publishing, Cham, 2019. https://doi.org/10.1007/978-3-030-15636-7_8.
- [18] LEHNERT, S.: *Music and Voice in German Newsreels of the 1950s/1960s*. In: *Proceedings of RE:SOUND 2019 – 8th International Conference on Media Art, Science, and Technology*, pp. 91–98, 2019. <http://dx.doi.org/10.14236/ewic/RESOUND19.15>
- [19] LLEIDA, E., A. ORTEGA, A. MIGUEL, V. BAZÁN-GIL, C. PÉREZ, M. GÓMEZ and A. DE PRADA.: *RTVE2020 Database Description*. 2020. catedrartve.unizar.es/reto2020/RTVE2020DB.pdf
- [20] NIETO FERRANDO, J.: *The Noticiari de Barcelona (1977-1980). The institutionalization of protest in the cinema of Spain's transition to democracy*, *Communication & Society* 32(1), pp. 279–292, 2019. <https://doi.org/10.15581/003.32.1.279-292>.
- [21] RAKAJ, D.: *An Analysis of Speech Acts in the Movie Hive*, *Theory and Practice in Language Studies* 13(11), pp. 2729–2733, 2023. <https://doi.org/10.17507/tpls.1311.01>
- [22] SCHWARZ, U.: *West German State Newsreels in the Period of the Economic Miracle 1950-1964: Gender as an Open Approach*. In: K. IMESCH, S. SCHADE, and S. SIEBER (eds.): *MedienAnalysen*, vol. 17, 1st edn., pp. 55–80, transcript Verlag, Bielefeld, 2016. <https://doi.org/10.14361/9783839429754-004>.
- [23] SEARLE, J. R.: *Expression and Meaning. Studies in the Theory of Speech Acts*. Cambridge University Press, Cambridge, 1979.
- [24] SIEBER, S.: 2016. *The Politics of Archives: Media, Power, and Identity*. In K. IMESCH, S. SCHADE, and S. SIEBER (eds.): *MedienAnalysen*, vol. 17, 1st edn., pp. 21–38, 1st edn. transcript Verlag, Bielefeld, 2016. <https://doi.org/10.14361/9783839429754-002>.

- [25] SUKIANI, N. K. and I. GUSTI WIRA HADIJAYA: *The Elements of Ethnography “Speaking” in Fist Fight Movies (2017)*. In: *Proceedings of the First International Seminar Social Science, Humanities and Education, ISSHE 2020, 25 November 2020*, Kendari, Southeast Sulawesi, Indonesia, 2020. <https://doi.org/10.4108/eai.25-11-2020.2306671>.
- [26] SOLBERG, P.E. and P. ORTIZ.: *The Norwegian Parliamentary Speech Corpus*. arXiv (arXiv:2201.10881). 2022. <https://doi.org/10.48550/ARXIV.2201.10881>.
- [27] THØGERSEN, J.: *The style and stylization of old news readings in Danish*. In: J. THØGERSEN, N. COUPLAND, and J. MORTENSEN (eds.): *Style, media and language ideologies* (Standard Language Ideology in Contemporary Europe 3), pp. 105–134, Novus Press, Oslo. 2016.
- [28] VASUDEVAN, R.: *Configurations of Partition: The Indian Newsreel Archive in the 1940s*. In: R. VASUDEVAN (ed.): *Media and the Constitution of the Political: South Asia and Beyond*. pp. 71–107, SAGE Publications, New Delhi, 2021. <https://doi.org/10.4135/978-93-5479-084-3>.
- [29] WIDAYANTI, S. R. and K. KUSTINAH: *Speech Act Analysis on Walt Disney Film Entitled „Frozen” (A Pragmatic Study)*. In: *Proceedings of the Fourth Prasasti International Seminar on Linguistics (Prasasti 2018)*. pp. 429–433. Atlantis Press, Surakarta, Indonesia, 2018. <https://doi.org/10.2991/prasasti-18.2018.79>
- [30] ZAND-VAKILI, E., A. FARD KASHANI, & F. TABANDEH: *The Analysis of Speech Events and Hymes’ SPEAKING Factors in the Comedy Television Series: “FRIENDS”*, *New Media and Mass Communication 2*, pp. 27–43, 2012.
- [31] Hungarian newsreel 400/4. 1931. <https://filmhiradokonline.hu/watch.php?id=7>
- [32] Hungarian newsreel 401/5. 1931. <https://filmhiradokonline.hu/watch.php?id=15>
- [33] Hungarian newsreel 404/5. 1931. <https://filmhiradokonline.hu/watch.php?id=37>
- [34] Hungarian newsreel 408/2. 1931. <https://filmhiradokonline.hu/watch.php?id=64>
- [35] Hungarian newsreel 412/2. 1932. <https://filmhiradokonline.hu/watch.php?id=82>
- [36] Hungarian newsreel 412/7. 1932. <https://filmhiradokonline.hu/watch.php?id=87>
- [37] Hungarian newsreel 413/8. 1932. <https://filmhiradokonline.hu/watch.php?id=96>
- [38] Hungarian newsreel 415/4. 1932. <https://filmhiradokonline.hu/watch.php?id=107>
- [39] Hungarian newsreel 416/8. 1932. <https://filmhiradokonline.hu/watch.php?id=119>
- [40] Hungarian newsreel 417/5. 1932. <https://filmhiradokonline.hu/watch.php?id=124>
- [41] Hungarian newsreel 418/1. 1932. <https://filmhiradokonline.hu/watch.php?id=128>
- [42] Hungarian newsreel 419/3. 1932. <https://filmhiradokonline.hu/watch.php?id=138>
- [43] Hungarian newsreel 420/6. 1932. <https://filmhiradokonline.hu/watch.php?id=147>
- [44] Hungarian newsreel 421/2. 1932. <https://filmhiradokonline.hu/watch.php?id=153>
- [45] Hungarian newsreel 428/4. 1932. <https://filmhiradokonline.hu/watch.php?id=198>
- [46] Hungarian newsreel 429/3. 1932. <https://filmhiradokonline.hu/watch.php?id=206>
- [47] Hungarian newsreel 430/3. 1932. <https://filmhiradokonline.hu/watch.php?id=214>
- [48] Hungarian newsreel 433/5. 1932. <https://filmhiradokonline.hu/watch.php?id=226>
- [49] Hungarian newsreel 437/3. 1932. <https://filmhiradokonline.hu/watch.php?id=250>
- [50] Hungarian newsreel 438/1. 1932. <https://filmhiradokonline.hu/watch.php?id=255>
- [51] Hungarian newsreel 440/4. 1932. <https://filmhiradokonline.hu/watch.php?id=270>
- [52] Hungarian newsreel 433/2. 1932. <https://filmhiradokonline.hu/watch.php?id=307>
- [53] Hungarian newsreel 449/1. 1932. <https://filmhiradokonline.hu/watch.php?id=333>
- [54] Hungarian newsreel 449/2. 1932. <https://filmhiradokonline.hu/watch.php?id=334>
- [55] Hungarian newsreel 452/4. 1932. <https://filmhiradokonline.hu/watch.php?id=363>
- [56] Hungarian newsreel 454/1. 1932. <https://filmhiradokonline.hu/watch.php?id=378>
- [57] Hungarian newsreel 454/2. 1932. <https://filmhiradokonline.hu/watch.php?id=379>
- [58] Hungarian newsreel 455/4. 1932. <https://filmhiradokonline.hu/watch.php?id=390>
- [59] Hungarian newsreel 457/2. 1932. <https://filmhiradokonline.hu/watch.php?id=406>
- [60] Hungarian newsreel 461/2. 1932. <https://filmhiradokonline.hu/watch.php?id=440>

A BRIEF HISTORY OF PITCH ANALYSIS

Philippe Martin

UFRL, Université Paris Cité, France

philippe.martin@utoronto.ca; philippe.martin@u-paris.fr

Abstract: Abandoning pitch events annotation based on their perception abilities, phoneticians and phonologists alike were gradually relying on acoustic analysis of speech, deemed to bring reliability and reproducibility to their measure of laryngeal frequency. This paper evokes the different steps implemented, made possible by the progress made in mechanical and electronic devices, operating first in the time domain then later in the frequency domain, starting from the kymograph to the most advanced signal analysis method operated on readily available computers.

1 Introduction

The measurement of pitch variation, usually represented by melodic curves in function of time, is closely linked to the development of advanced mechanical then electronic systems. Next to the abundance of research on the phonetic and linguistic aspects on prosody, there is little work on the history of the equipment associated with this field, with the possible exception of Odette Mettas monograph [1], a chapter in Pierre Léon and Philippe Martin's book [2] and a paper by Martin dedicated to the subject published in 2011 [3]. However, a large number of papers have been devoted to the signal analysis aspects of pitch tracking (Hess [4]), which were largely ignored by prosodists at large.

2 Looking at the speech waveform: the kymograph

The kymograph was invented in 1847 by Carl Ludwig. Originally, it was used to record muscle movements, changes in blood pressure and other physiological phenomena via specialized sensors. The kymograph has then been used in experimental phonetics research to record changes in air pressure in the nasal, buccal and labial cavities, as well as variations in pressure at the lips, and thus, in a rudimentary way, the speech signal (Teston [5]). The kymograph records used a variety of media, which has to move in function of time as regularly as possible (from lampblack-coated paper, graph paper, photosensitive film, cinematographic film, etc.).

As early as 1897, in his thesis, Rousselot used kymographic recordings for his dialectological research, and studied ways of obtaining characteristic graphic representations of speech (Figure 1).

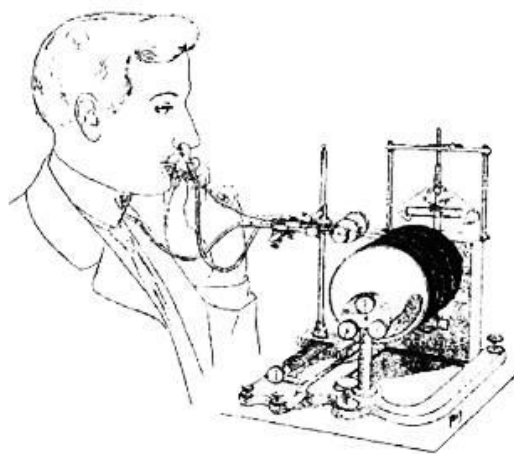


Figure 1. Rousselot kymograph (Rousselot et Laclotte [6]). A stylus traduces the variation of mouth air pressure on a lampblack-coated rotating cylinder

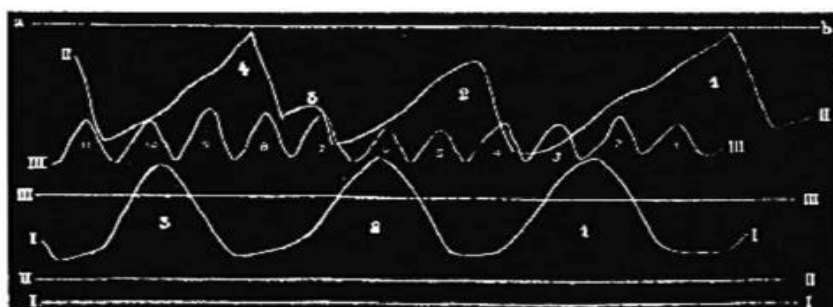


Figure 2. Kymographic recording allowing measurement of successive laryngeal periods by visual inspection (Vierordt et Ludwig 1855, from Panconcelli Calzia [7]).

Pressure variations picked up by an acoustic horn (later by a microphone), are recorded by a metal stylus, whose inertia has to be as low as possible, so that the bandwidth achieved by the entire mechanical system is satisfactory, and the most important vibrations (those corresponding to mouth pressure) recorded as faithfully as possible (Figure 2).

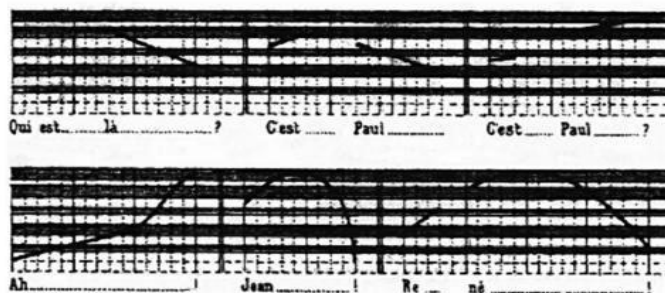


Figure 3. An example of a melodic curve obtained from a kymographic trace (Marichelle [8]).

Knowing the speed of movement of the inscribable medium, successive time intervals between two equivalent moments of the laryngeal cycle, defined as the inverse of the laryngeal period, can be estimated. In general, these moments correspond to the maxima (or minima) of the recorded curves, or to zero crossings, which a trained operator can identify by visual inspection in an expected time interval. From the measurement of a succession of periods, the corresponding instantaneous laryngeal frequency can be plotted over time. This curve is often referred to as the melodic curve. In France, the first measurements were published by Hector Marichelle in 1897 [8] (Figure 3).

Later, with the development of electronic devices, the stylus was replaced by an ink jet (mingograph), then by an ultraviolet light beam deflected in time with the vibrations by a small half-mirror printing on photosensitive paper (Honeywell UV recorder).

Precise calculation of laryngeal frequency can be carried out either by evaluating instantaneous variations, by measuring each period, or by calculating an average over n periods (e.g. $n = 10$). In the first case, good temporal definition is achieved at the expense of accuracy (due to measurement error caused by line thickness); in the second case, accuracy is n times greater, but averaging smoothes out instantaneous variations. (Figure 4). Again, this presupposes that there is only one identified peak within the expected period, which implies that the range of period values should be estimated in advance.

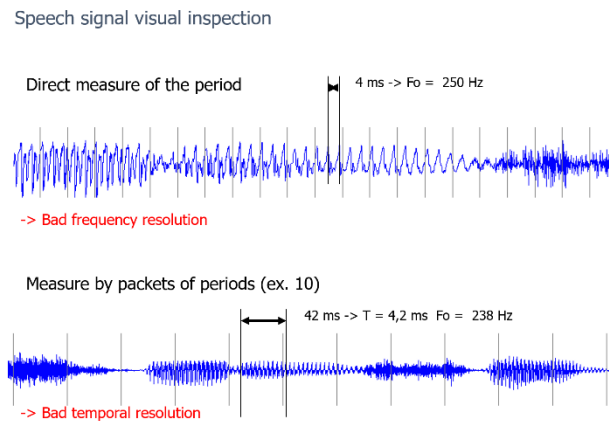


Figure 4. An example of plots illustrating the two types of measurement: period by period (above) and by packet of ten periods.

The whole reliability of the system relies on the user ability to effectively distinguish the successive periods from this representation of the speech wave. One noticeable and well-known problem pertains to the vowel [u], where the amplitude of the second harmonic can give the illusion that the period is actually half of what it is really. Another case pertains to (some) creaky voices, where the apparent period is actually the double of the laryngeal frequency, confusing even the Fourier transform in its harmonic analysis.

In fact, McKinney [9] gives the condition so that the first harmonic appears with a larger amplitude in each cycle, and therefore avoid possible confusion for human and machine operators alike: the amplitude of the first harmonic must be higher than the sum of amplitudes of all other harmonics (i.e. from the 2nd harmonic).

3 Time domain pitch trackers

As early as 1935, collaboration between phoneticians and electronic engineers led to the development of a large number of devices for direct measurement of laryngeal frequency, based on the detection of periodic peak patterns in the signal. The principle of manual measurement was reproduced electronically by automatic peak detection of the signal collected by a microphone, amplified and processed by filtering. A detailed analysis of the problems involved in measuring laryngeal frequency, either directly using a laryngograph or indirectly by processing the speech signal, was carried out by McKinney [9].

Instead of estimating F_0 from the acoustic signal, some detectors (laryngoscopes and laryngographs) have used physiological sensors of various kinds: measurement of impedance variations in the speaker's larynx, use of electrodes (Fourcin [11]), and so on. Of course, this type of approach requires the presence of the speakers, but it has the advantage of capturing signals with very low background noise.

3.1 Manually adjustable filters

The first so-called melody detectors, using radio tubes, date back to the 1930s (Hunt 1935 [11], Obata and Kobayashi 1937 [12], Grützmacher and Lottermoser 1937 [13]). The eventual low amplitude of the fundamental compared with the second harmonic often defeats this

strategy, as two peaks of close amplitude (or several zero crossings) can occur in the same signal period.

To improve performance, low-pass filters have often been used, providing approximate inverse filtering of the vocal tract to recover the laryngeal source signal. At first, filters were manually adjustable (e.g., a device marketed by Frøkjær-Jensen Electronics and designed by Gunnar Fant, see Fant 1958 [14]) (Figure 5), and required different filters according to the speaker's vocal range. Later, self-switching filters appeared, enabling the measurement of rapid variations in pitch (Pitch Computer, Frøkjær-Jensen Electronics, designed by Martin [15]).



Figure 5. Melody detector with manually switched filters (Fundamental Frequency Meter, Frøkjær-Jensen Electronics).

To overcome the variability of the speech signal, a number of designs have been tested in an attempt to improve pitch detection [4].

In France, a number of designs were developed, made more affordable by the widespread use of semiconductors, which gradually replaced radio tubes. These set-ups became relatively simple to reproduce, taking into account the time-domain properties of the speech signal. The Grenoble research teams put into practice concepts developed jointly by phoneticians and electronics engineers (Carré et al. 1963 [17], Boë and Rakotofiringa 1971 [18]).

Later, at the *Institut de phonétique* in Paris, Bernard Gautheron developed a melody analyzer that also operated in the time domain, while Bernard Teston collaborated with Mario Rossi at the Phonetic Institute in Aix-en-Provence to develop a melody detection system (Teston 1972, [19] Teston and Rossi 1977 [20]).

Time-domain analysis methods present a number of difficulties due to the highly variable intra- and interlocutor characteristics (e.g., rapid laryngeal frequency changes), and to recording conditions that are not always those of a sound proof room. The most frequent faults with analyzers are poor measurement of the fundamental, whose value may be doubled (octave jump) or halved, as well as significant errors at unvoiced/voiced transitions in the analyzed signal. Peak detection measurement is hampered by rapid amplitude and frequency variations and the high intensity of the second harmonic.

Some recordings were made with automatic volume control, which is particularly disruptive to the analysis (whereas a naïve approach to the problem might lead to hope for improvement, since the signal is in a sense normalized). The hysteresis inherent in electronic zero-crossing detection makes the period measure dependent on the change in signal amplitude.

3.2 Automatically adjustable filters

A first implementation in the Experimental Phonetics laboratory created by P. Léon in 1966, an automatic filtering system has been implemented using a PDP8/I, an expensive machine with 8 kB of 12-bit memory...delivering a pitch curve in real time on the remanent screen of a Tektronix oscilloscope. A camera allowed the display of a pitch curve on a large

TV screen. The whole system worked in real-time on a large set of male or female voice from a microphone input, which allowed Ph. Martin to analyze quickly a large number of examples, which would have taken many days if not months at the time, as most research on intonation was carried out on spectrographs (e.g., Delattre [10]). This would lead eventually to propose an intonation model featuring the contrast of melodic slope in French.

This design was improved further by Martin (1977 [15]) using a bank of filters to sufficiently eliminate higher harmonics over a wide frequency range and, after compensating for the different phase shifts of each filter channel, performing a symmetrical zero-crossing measurement (positive and negative values measured separately). A microprocessor (Intel 8080) placed at the filter output performs the period measurements, as well as the necessary corrections and smoothing.

Since cycle-by-cycle detection almost always gives rise to errors under real conditions (poor detection of the second harmonic, false estimation of voiced/unvoiced segments, skipping of an amplitude peak), electronic and then computerized processing were designed very early on to correct them. This system has been marketed first by the Danish firm F-J Electronics under the name of “Pitch Computer”, and later by Voice Identification in New Jersey.

4. Frequency domain pitch trackers

Instead of tracking the speech signal successive peaks, an harmonic analysis of the speech signal by a Fourier transform appears as a more appropriate solution. However, some drawbacks had to be considered, such as the lost of resolution on the time scale (due to the time window implied to obtain a satisfactory frequency resolution), and lack of fast electronic analog designs at the time. Later, the advent of affordable personal computers allowed the implementation of complex software routines using all the information provided by a spectrogram and not just the fundamental frequency.

4.1 From the flame spectrograph to the sonagraph

Spectral analysis of speech dates back to the 19th century. The flame spectrograph designed by Kœnig in the early 1860s, based on the work of Helmholtz, was used for the first acoustic analyses of speech.



Figure 6. Kay Elemetrics Spectrograph « Sonagraph » (photo Kay Elemetrics)

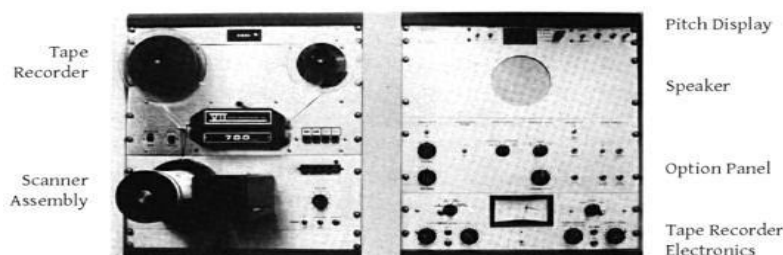


Figure 7. Voice Identification Spectrograph (photo Voice Identification).

Later, the development of tube and then solid-state electronics made it possible to mass-produce devices that were easier to use and, although quite expensive at the time, within the reach of experimental phonetics laboratories.

The two best-known manufacturers, who maintained a virtual monopoly on these instruments for a long time, were the Americans Kay Elemetrics and Voice Identification, whose workshops were both located in New Jersey, just a few miles from each other. The Kay spectrograph, smaller and less expensive, at least in its best-selling version, uses a rotating magnetic disk as a recording and playback medium.

The more expensive Voice ID spectrograph operates directly from a magnetic tape on which the speech to be analyzed has been pre-recorded. In both cases, however, analyses are performed in time windows of exactly 2.4 seconds.

The principle of spectrogram tracing is the same for both devices: a stylus is subjected to an electrical voltage proportional to the intensity of the speech signal filtered by a variable-frequency heterodyne bandpass filter. This heats up and blackens a sheet of heat-sensitive paper wrapped around the drum. The stylus is driven by a worm screw along an axis, whose displacement coincides with the sliding of the analysis frequency of the bandpass filter. The position of the analysis drum is vertical for Kay Elemetrics and horizontal for Voice Identification, as shown in figures 6 and 7.

Measure the fundamental frequency from a narrow-band spectrogram, by measuring a visually identified high harmonic (e.g. the 10th).

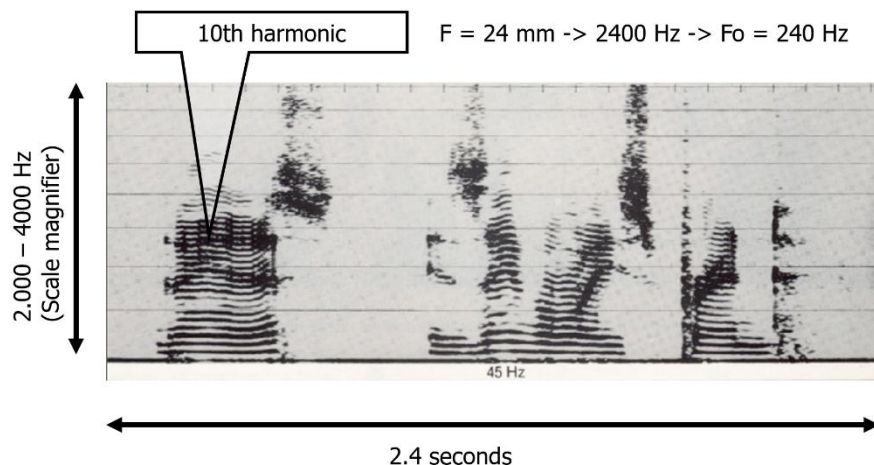


Figure 8. Measure the fundamental frequency from a narrow-band spectrogram, by measuring a visually identified high harmonic (e.g. the 10th).

One revolution of the drum corresponds to 2.4 seconds of speech, and the stylus travels along the worm screw in an average of almost 3 minutes. Two bandwidths (300 hertz and 45 hertz) enable different resolutions to be achieved, traditionally referred to as broadband and narrowband. The wide band is suitable for measuring formants, and the narrow band for harmonics, and therefore voice pitch. In retrospect, we can admire the patience and determination of the phoneticians of the time, analyzing 10-minute corpora (at least 90 times real time was required).

In the narrow band, harmonics appear as relatively thick lines (corresponding to 45 hertz on the frequency scale). To improve measurements, frequency values are often taken from the 10th harmonic (if visible). For various material reasons - imperfections in electrical resistors (potentiometers), overheating of tube and then transistor amplifiers — the frequency scale is often non-linear (and even fluctuating during operation). In such cases, a calibration signal

(e.g. with harmonics spaced at 500 hertz) must be used to better estimate the frequency of a particular harmonic.

In any case, while the general shape of the melodic curve can be reasonably observed using narrow-band spectrographic analysis, the accuracy of the measurement is quite low. If we assume a length measurement error of ± 0.5 mm over 24 mm, corresponding to 2400 hertz for example, the resulting frequency measurement error is ± 48 hertz/10, i.e. ± 5 hertz approximately.

5. Spectral algorithms

Rapid improvement of computer speed, the incorporation of dedicated multiplier sections in core chips lead to the development of software performing real-time Fourier transform (using the Fast Fourier Transform or FFT). One of the first dedicated system operating in real-time, designed by Ph. Martin, was presented in the exhibition *Les Immatériaux* in the *Centre Beaubourg* in Paris in 1985 (Color real-time spectrograph).

Entire books have been devoted to F0 measurement circuits and algorithms [4]. The most recent, implemented in software such as Praat or WinPitch, use all the spectral information of neighboring sounds, generally obtained by Fourier transform or autocorrelation. These include cepstrum, which detects the fundamental frequency by spectral analysis of the logarithm of the signal's amplitude spectrum, and spectral comb, which searches for the fundamental frequency by intercorrelation with a variable-interval comb function. (Martin 1981 [22]).

These methods enable reliable measurement even in the case of signals where the fundamental is absent (as in old analog telephone recordings) or which are noisy, provided that the spectral distribution of the noise is sufficiently different from that of the speech signal. In principle, these methods use all available spectral information.

The use of spectral methods limits temporal resolution, less suitable for very low-pitched voices (with a fundamental at 80 hertz, for example) than for high-pitched voices (for which F0 is of the order of 250 hertz). Indeed, to obtain sufficient frequency resolution for a low fundamental frequency, a spectral analysis window of at least 60 milliseconds is required. Details of rapid variations in the fundamental frequency are therefore necessarily lost, whereas with a higher voice, a window of 30 milliseconds or less may be appropriate, and a better temporal resolution achieved. A variable-window system can then be implemented to automatically optimize temporal resolution.

The spectacular reduction in computation times for spectral analysis has encouraged several researchers to pursue harmonic separation in the case of multi-source recordings. In France, we can cite Martin (2000 [23], 2008 [24]) with the “spectral brush” method and Liénard, Signol and Barras (2007 [25]) with the inverse comb calculation.

Pitch analysis software

The generalization of personal computers operation on MacOs or Windows allowed the emergence of a rather large number of software packages with at least one of their functions devoted to voice pitch tracking and fundamental frequency display. Among the most popular, there is of course Praat [26], with its pitch tracking based on autocorrelation. One of the first commercialized available packages is Signalize (on Mac, 1990 [27]), developed by Eric Keller, using Autocorrelation and FFT Martin’ Spectral Comb [22]. Other systems which were less known are PitchWorks [28], Speech Analyzer [29] Computerized Speech Lab [30], xassp [31].

In its latest version, WinPitch for instance, incorporated Whisper routines, fine-tuned to transcribe spontaneous speech “errors” (i.e., repetition, abandons, hesitations,

reformulations...) accurately. The software embeds automatic API segmentation, as well as Text-to-Speech routines in more than 50 languages. Pertaining to fundamental analysis, it allows the user to visually inspect the validity of segments of F0 curve, by comparison with an underlying narrow band spectrogram whose frequency scale is automatically aligned on the F0 scale. Erroneous F0 segments can be graphically corrected, allowing even to classify prosodic events according to a user defined model (ToBI, Glissando, etc.), and display automatically the prosodic structure as indicated by the transcribed prosodic events. This particular feature addresses the current problems of pitch analysis, leading researchers, faced with unreliable pitch curves, to rely on sound proof recording to obtain reliable data.

In practice, fundamental frequency tracking presents numerous specific problems in real-life:

- a) possible absence of first harmonic, due to filtering, wrong setup of microphone, historical recordings...
- b) low signal level, sometimes below the recording channel noise level, making any pitch tracking routine to operate satisfactory
- c) voice or other sound sources overlapping, making most routines based on harmonic detection failed
- d) high level background noise

The problems can be sometimes alleviated by deep learning approaches, where interfering signals can be trained to be separated from speech. But as requiring training of somewhat large number of samples, their use is practically limited to large corpora.

However, for most applications, users can be easily trained to separate visually first (or second) harmonic of specific speech segments if displayed with the same frequency scale than the partially erratic pitch curve. With designated graphic commands, the user can then select and manually enter on screen the appropriate pitch segments (Figure 9). This allows for accurate transcription of prosodic events in degraded recordings, often found in spontaneous speech data.

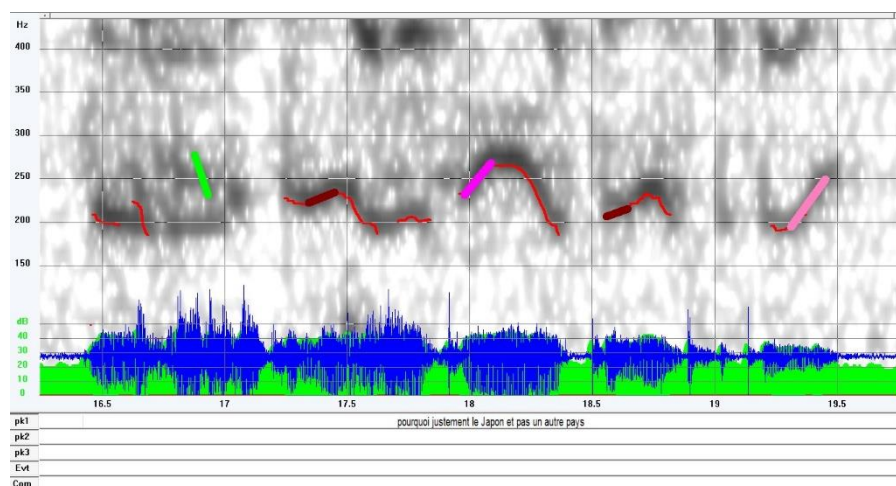


Figure 9. Simultaneous fundamental frequency and narrow band spectrogram on the same frequency scale, allowing visual verification of the Fo values by comparison with the spectrogram first (or second) harmonic, and their correction with integrated graphic commands (WinPitch).

References

- [1] METTAS, O.: *Les techniques de la phonétique instrumentale et l'intonation*, Bruxelles, Presses universitaires de Bruxelles et Paris, Maloine, 1091.
- [2] LÉON, P. et MARTIN, Ph.: *Prolégomènes à l'étude des structures intonatives*, Studia Phonetica 2, Didier : Près, 226 p. 1969.
- [3] MARTIN, Ph.: Éléments d'histoire de l'analyse de la fréquence laryngienne, in *Un siècle de phonétique expérimentale*, L.-J. BOË et C.-E. VILAIN éd., ENS Éditions, pp. 319-329, 2010.
- [4] HESS, W.: *Pitch Determination of Speech Signals*, Berlin: Springer-Verlag, 698 p., 1983.
- [5] TESTON, B.: Le rôle d'Étienne-Jules Marey dans l'émergence de la phonétique expérimentale, in *Un siècle de phonétique expérimentale*, L.-J. BOË et C.-E. VILAIN éd., ENS Éditions, pp. 59-89, 2010.
- [6] ROUSSELOT P.-J., LACLOTTE F.: *Précis de prononciation française*, Paris, H. Welter, 1902.
- [7] PANCONCELLI-CALZIA, G.: Earlier History of Phonetics, *Manual of Phonetics*, Louise Kaiser éd., Amsterdam: North Holland, pp. 3-17, 1957.
- [8] MARICHELLE H. V.: *Phonétique expérimentale. La parole d'après le tracé du phonographe*, Paris, Delagrave, 198 pages, 1897.
- [9] MCKINNEY N. P.: *Laryngeal Frequency Analysis for Linguistic Research*, Ann Arbor, University of Michigan Communication Science Laboratory, 1965.
- [10] DELATTRE P.: Les dix intonations de base du français, *French Review* 40, pp. 1-14, 1966.
- [11] FOURCIN A.: Laryngographic analysis of vocal fold vibrations, *Ventilatory and Phonatory Control Systems: an International Symposium (Royal College of Surgeons of England, 1972)*, Barry Wyke ed., London: Oxford University Press, pp. 315-333, 1972.
- [12] HUNT F. V.: A direct-reading frequency meter suitable for high-speed recording », *The Review of Scientific Instruments*, 6, pp. 43-46, 1935.
- [13] OBATA J., KOBAYASHI R.: A direct-reading pitch recorder and its applications to music and speech », *Journal of the Acoustical Society of America*, 9 (2), pp. 156-161, 1937.
- [14] GRÜTZMACHER M., LOTTERMOSER W.: Über ein Verfahren zur trägheitsfreien Aufzeichnung von Melodiekurven, *Akustische Zeitfrift*, 2, pp. 242-248, 1937.
- [15] FANT G.: Modern instruments and methods for acoustic studies of speech, *Acta Polytechnica Scandinavica*, n° 1, pp. 1-106, 1958.
- [16] MARTIN Ph.: Un analyseur-visualiseur de mélodie à microprocesseur, *Actes des VIII^e Journées d'étude sur la parole, Aix-en-Provence, 1977*, Lannion, GALF, pp. 95-102, 1977.
- [17] COWAN J. M.: Graphical representation of perceived speech », *Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki, 1961*, Antti Sovijäri et Pentti Aalto ed., La Hague, Mouton, pp. 569, 1962.
- [18] CARRÉ R., LANCIA R., PAILLÉ J., GSELL R.: Étude et réalisation d'un détecteur de mélodie pour analyser la parole, *L'Onde électrique*, 434, pp. 556-562, 1963.
- [19] BOË L.-J., RAKOTOFIRINGA H.: Exigences, réalisation et limites d'un appareillage destiné à l'étude de l'intensité et de la hauteur d'un signal acoustique, *Revue d'acoustique*, 4, pp. 104-113, 1971.

- [20] TESTON B.: Description d'un système de détection de l'intensité et de la mélodie de la voix, *Travaux de l'Institut de phonétique d'Aix-en-Provence*, 1, pp. 129-145, 1972.
- [21] TESTON B., ROSSI M.: Un système de détection automatique de la fréquence fondamentale et de l'intensité de la parole, *Actes des VIII^e Journées d'étude sur la parole, Aix-en-Provence, juin 1977*, Lannion, GALF, pp. 111-117, 1977.
- [22] MARTIN Ph.: Mesure de la fréquence fondamentale par intercorrélation avec une fonction peigne », *Actes des XII^e Journées d'étude sur la parole, Montréal, mai 1981*, Montréal, direction des communications de l'université de Montréal.
- [23] MARTIN Ph.: Peigne et brosse pour F_0 : mesure de la fréquence fondamentale par alignement de spectres séquentiels », *Actes des XXIII^e Journées d'étude sur la parole, Aussois, juin 2000*, Grenoble, Institut de la communication parlée, pp. 245-248, 2000.
- [24] MARTIN Ph.: Crosscorrelation of adjacent spectra enhances fundamental frequency estimation », *Proceedings of Interspeech 2008 incorporating SST 2008, Brisbane, 2008*, Rudle Mall, Causal Productions, pp. 1485-1488, 2008.
- [25] LIÉNARD J.-S., SIGNOL F., BARRAS C.: Speech fundamental frequency estimation using the alternate comb, *Interspeech 2007. Proceedings of the VIIIth Annual Conference of the International Speech Communication Association, Antwerp, 2007*, Red Hook, NY, Curran, 3, pp. 46-50, 2007.
- [26] Praat <https://www.praat.org>
- [27] Keller, E. *Signalize™: Signal analysis for Speech and Music. User's Manual*. Rosemère, Québec: InfoSignal Inc. 1990,
https://joaquimllisterri.cat/phonetics/fon_anal_acus/herram_anal_acus.html#Signalize_1
- [28] PitchWorks, Scicon R&D <https://sciconrd.com/pcquirerx.aspx>
- [29] Speech Analyzer, SIL International <https://software.sil.org/speech-analyzer/>
- [30] CSL, Computerized Speech Lab, Pentax Medical
<https://web.archive.org/web/20240719021705/https://www.pentaxmedical.com/pentax/en/99/1/Computerized-Speech-Lab-CSL>
- [31] xassp, IPDS Institut für Phonetik und digitale Sprachverarbeitung, Christian-Albrechts-Universität, Kiel <https://www.ipds.uni-kiel.de/forschung/xassp.en.html>

TRACING PITCH ON THE GRID PAPER: FU LIU AND THE LOGARITHMIC SCALE PARADIGM BEFORE THE DIGITAL ERA

Jingyi Sun, Jiāyīn Gāo, Annie Rialland

Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle), France

jingyi.sun@sorbonne-nouvelle.fr

Abstract: This paper reconstructs the theoretical rationale and technical implementation of a logarithmic scale paradigm for pitch analysis devised by Chinese phonetician Fu Liu (刘复) in the 1920s. Faced with the challenge of representing tones, which requires both fine-grained measurement and the visualization of psychoacoustically motivated pitch contours, Liu designed a system based on logarithmic spacing and twelve-tone equal temperament. Unlike his European and American contemporaries such as Daniel Jones, Roussetot, and Scripture, who relied on perceptual judgment, proportional musical scales, or direct frequency tracing, Liu sought a reproducible framework for tonal analysis grounded in numerical precision. Drawing on original documents and comparative analysis, this study situates Liu’s work within the broader context of early 20th-century phonetics. The logarithmic scale is now widely used in modern acoustic analysis tools, but Liu’s accurate implementation of the scale was achieved manually before our digital era, which should be considered as a pioneering contribution. Through the lens of Liu’s legacy — often overlooked — we show the importance of cross-cultural and cross-disciplinary exchange in shaping the early development of phonetic sciences.

1 Introduction

At the turn of the 20th century, experimental phonetics was taking shape as a scientific discipline in Europe and North America. Researchers aimed to measure physical properties of speech such as pitch, intensity, and duration, not just to document language variation, but to uncover the underlying physiological and acoustic mechanisms. Pitch, in particular, posed technical and theoretical challenges. Daniel Jones proposed a perception-based notation system relying on trained musical judgment [7]; Abbé Roussetot used acoustic recordings and proportional scales based on just intervals [20]; Wheeler Scripture advocated direct tracing of frequency curves in hertz [22]. While these methods were influential in their time, we have found no evidence that the scholars employed a logarithmic scale or developed a fully standardized system for analyzing and visualizing pitch in tonal or intonational contours.

Within this context, the approach developed by Fu Liu presents a noteworthy departure [26, 14]. Drawing on both Chinese and Western intellectual traditions [15], Liu constructed a logarithmic pitch scale system grounded in twelve-tone equal temperament and anchored to a 435 Hz reference frequency. His method combined lookup tables and semitone-divided grid paper to enable manual yet systematic visualization of pitch movement across speakers and dialects. Although devised without access to modern computing, Liu’s paradigm anticipates many of the core principles behind later software tools such as Praat [3]. This paper reconstructs Liu’s logarithmic scale paradigm in both its theoretical and technical dimensions and considers its significance in the broader trajectory of phonetic research.

2 Early 20th-Century Attempts at Pitch Measurement

2.1 Daniel Jones: Perception-Based Notation

Among the early pioneers of experimental phonetics, Daniel Jones proposed a method of pitch analysis that relied primarily on auditory perception. In his 1909 work *Intonation Curves* [7], Jones rejected acoustic measurement procedures as too laborious to be of practical use, arguing that “no one has ever yet analyzed a text of sufficient length to be of practical value to language students.” (*Jones 1909*, Introduction, p. IV)

Instead, he proposed a method based on the phonetician’s immediate auditory impression, which involved playing a phonograph recording and briefly lifting the needle to isolate a single moment of sound. The listener would then rely on memory to identify the quality and, if the sound was voiced, its pitch. (*Jones 1909*, Introduction, p. V)

This method required the phonetician to have a well-trained musical ear, capable of judging absolute pitch from brief auditory impressions. Using this approach, Jones was able to provide notations for large corpora. An example of the intonational notation of the beginning of *La Samaritaine* by Sarah Bernhardt is shown in Figure 1¹.

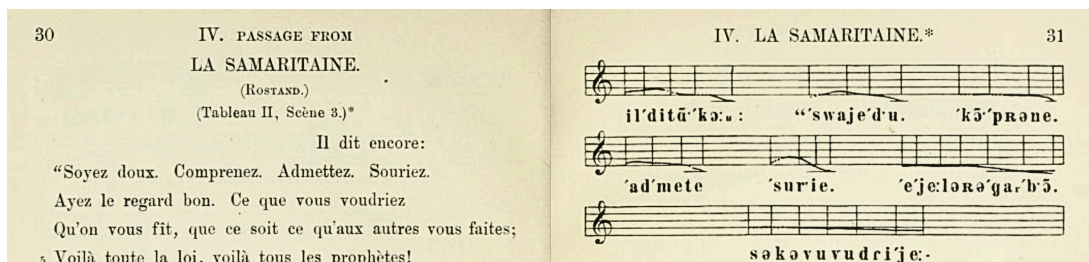


Figure 1. Daniel Jones’s intonational notation of *La Samaritaine* as performed by Sarah Bernhardt. *Jones 1909*, p. 30-31.

The choice to use conventional musical notation was driven by its familiarity to students studying foreign languages such as English, German, or French, which is the target audience for this book. In 2020, Ashby conducted a comparative analysis between a modern pitch analysis of Jones’s corpora and Jones’s original work [1]. His findings revealed a significant convergence between the two, thereby validating Jones’s methodology retrospectively, at least in the context of studying English intonation.

However, Fu Liu remained unconvinced of the method. While at University College London, he observed Stephen Jones (an assistant to Daniel Jones) and an Indian student applying the technique to Sanskrit tone studies. Although he did not participate in these experiments, he later commented on them in his 1929 Chinese treatise *Shēngdiào zhī pànduàn jí shēngdiào pànduàn chǐ zhī zhìzào yǔ yòngfǎ* [12], written after his return to China. He described the method as complex, imprecise, and overly reliant on subjective impression, arguing that many pitch determinations were made without sufficient perceptual anchoring and thus lacked consistency and reliability. Yuen Ren Chao (趙元任), who had also experimented with the method, reached a similar conclusion: that the fleeting auditory cues it relied on were inadequate for capturing full pitch trajectories [12].

Liu’s ideas about how pitch should be analyzed were much closer to those of Rousselot and Scripture, both of whom had worked at Rousselot’s laboratory and placed great importance on acoustic evidence. Liu shared Rousselot and Scripture’s view that acoustic data should form the empirical basis of pitch analysis — an aspect notably absent

¹All historical figures cited in this paper are available in digitized form online. For the reader’s convenience, hyperlinks are included in the captions.

from the work of Jones, who emphasized perceptual and physiological data, particularly in *An Outline of English Phonetics* (1922) [8]. While Jones successfully published this handbook as a practical guide, Liu and Chao considered the approach too individualistic and insufficiently replicable. Moreover, even if the method might suffice for describing intonation in stress-based languages like English, it was inadequate for tonal languages such as Mandarin, where pitch contour is lexically contrastive and requires much greater acoustic resolution.

2.2 Rousselot: Acoustic Recording and the Natural Scale

In his 1891 doctoral thesis *Les modifications phonétiques du langage étudiées dans le patois d'une famille de Cellesrouin*, Abbé Rousselot devoted a chapter to pitch measurement (“Mesure de la hauteur des sons”) [19]. As a native speaker of the Cellesrouin dialect, he recorded his own speech using an electric laryngeal probe and a kymograph. These devices enabled him to trace glottal vibrations and determine the frequency of each vibratory cycle.

He then plotted his findings on a musical staff, using a proportional scale constructed from simple integer ratios. The reference note *la* (i.e., *A* in the Anglo-German system)² was set at 426.26 Hz, reflecting the tuning standard of his instrument-maker, Koenig, and slightly influencing the frequency values he annotated. An example of his pitch transcription, based on the beginning of the tale *Le petit Poucet*, is shown in Figure 2. Rousselot provided pitch notations not only for isolated words and full sentences, but

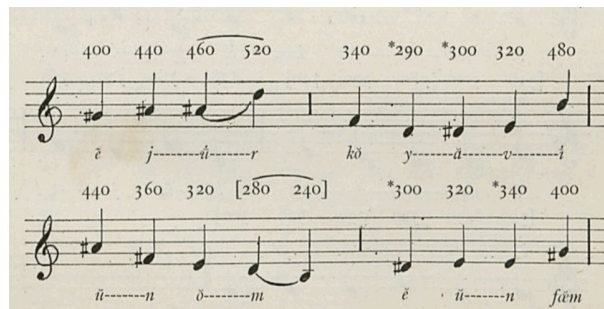


Figure 2. Rousselot’s musical transcription of the tale opening *Le petit Poucet*. Rousselot 1891, p. 133

also for entire folktale openings. He compared his results with transcriptions made by a professional musician and observed a notable correspondence between them.

In *Principes de phonétique expérimentale* [21], Rousselot formalized the system further by elaborating the theoretical basis of his proportional scale, which he called the *gamme naturelle*³. Unlike equal temperament, this scale was derived from just intonation ratios such as 9:8, 4:3, and 5:4. He further extended these ratios to calculate all semitones and provided detailed correspondence tables between frequencies and semitone positions, covering nine octaves (ranging from 16.3125 Hz to 7830 Hz), as shown in Figure 3. This version of the scale was anchored to 435 Hz, consistent with the French diapason standard of the time. Rousselot provided frequency-to-semitone conversion tables covering nine octaves, enabling systematic analysis across a wide pitch range. For graphical represen-

²Throughout this paper, solfège pitch notation (e.g., la_3 , ut_0) follows the French system commonly used in early phonetic studies. Where relevant, equivalent pitch names from the international (Anglo-German) letter system (e.g., A_4 , C_1) are provided in parentheses for clarity.

³Rousselot’s use of the term *gamme naturelle* refers to a just-intonation scale based on harmonic proportions, not to a diatonic scale in the sense of a seven-note stepwise sequence such as C–D–E–F–G–A–B.

NOTES	OCTAVES								
	Ut ₋₉	Ut ₋₈	Ut ₋₇	Ut ₋₆	Ut ₋₅	Ut ₋₄	Ut ₋₃	Ut ₋₂	Ut ₋₁
Ut...	16,3125	32,625	65,25	130,5	261	522	1044	2088	4176
Ré...	18,3515625	36,703125	73,40625	146,8125	293,625	587,25	1174,5	2349	4698
Mi...	20,390625	40,78125	81,5625	163,125	326,25	652,5	1305	2610	5220
Fa....	23,75	47,5	95	190	380	760	1520	3040	6080
Sol...	24,46875	48,9375	97,875	195,75	391,5	783	1566	3132	6264
La....	27,1875	54,375	108,75	217,5	435	870	1740	3480	6960
Si ♭...	29,3625	58,725	117,45	234,9	469,8	939,6	1879,2	3758,4	7516,8
Si.....	30,5859375	61,171875	122,34375	244,6875	489,375	978,75	1957,5	3915	7830
Ut ♯...	16,9921875	33,984375	67,96875	135,9375	271,875	543,75	1087,5	2175	4350
Ré ♯...	19,1162109375	38,232421875	76,46484375	152,9296875	305,859375	611,71875	1223,4375	2446,875	4893,75
Fa ♯...	22,65625	45,3125	90,625	181,25	362,5	725	1450	2900	5800
Sol ♯...	25,48828125	50,9765625	101,953125	203,90625	407,8125	815,625	1631,25	3262,5	6525
La ♯...	28,5203125	56,640625	113,28125	226,5625	453,125	906,25	1812,5	3625	7250

Figure 3. Frequency-to-semitone table across nine octaves in Rousselot’s proportional system. *Rousselot 1897-1901*, p. 12

tation, he allocated equal horizontal spacing to each semitone, diverging from traditional musical notation. This is shown in Figure 4. The frequency curves generated by these

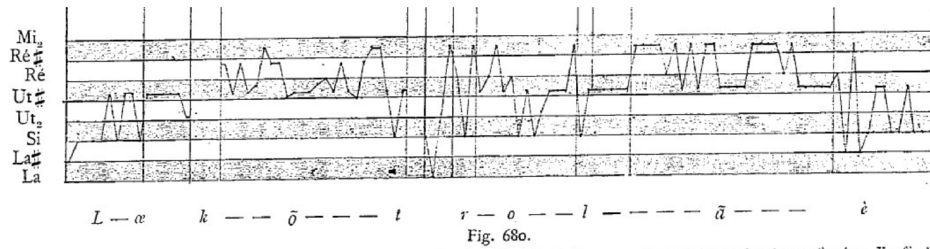


Figure 4. Graphical design of Rousselot’s system with uniform semitone spacing, diverging from traditional music notation. *Rousselot 1897-1901*, p. 1004

recordings appeared jagged and uneven, requiring substantial manual smoothing, a difficulty also noted by Fu Liu in his later writings. While Rousselot adopted the terminology of twelve semitones per octave, these intervals were not equally spaced logarithmically. Rather, they were constructed from harmonic ratios in line with the principles of just intonation. This distinguishes his proportional scale from the equal-tempered system later adopted by Fu Liu.

2.3 Wheeler Scripture: Hertz-Based Acoustic Curves

Wheeler Scripture maintained a close professional relationship with Rousselot and visited his laboratory in Paris, where he conducted experiments using various instruments and recording techniques. Sharing Rousselot’s emphasis on physical measurement, Scripture strongly advocated grounding intonational research in acoustic analysis.

In his book *Elements of Experimental Phonetics* (1902) [22], he argued that the constantly changing nature of speech pitch made it virtually impossible to capture accurate melodic contours by ear alone. Instead, he promoted the use of phonographic devices to trace pitch variation continuously over time, a principle especially relevant to tonal languages like Chinese.

Scripture’s empirical work relied exclusively on frequency-based acoustic measurements, expressed in hertz. His work ranged from single sentences to various speech forms, including a performance of *Rip Van Winkle’s Toast* by actor John Jefferson and a public address by Senator Depew. An illustrative example appears in Figure 5, depicting

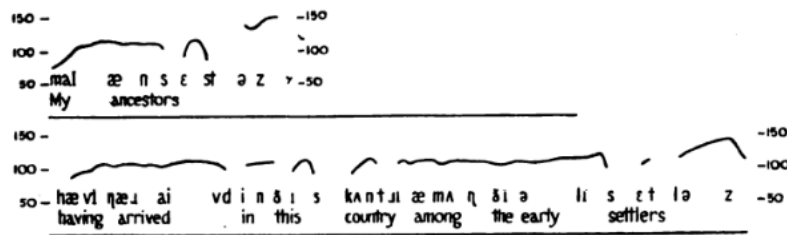


Figure 5. Scripture’s pitch tracing of the beginning of Senator Depew’s public address, measured in hertz. *Scripture 1923*, p. 11.

the pitch contour of the beginning of senator Depew’s address taken from *The Study of English Speech by New Methods of Phonetic Investigation* [23].

Unlike Rousselot, Scripture made no attempt to systematize pitch using musical frameworks, whether proportional or logarithmic. His contribution was instead rooted in the direct acoustic visualization of pitch movement.

2.4 Jean Poirot: Interval Percentages and Cross-Linguistic Inquiry

Before moving to Paris, where he was appointed to the Institut de Phonétique, Jean Poirot held a chair in phonetics at the University of Helsinki. During his time there, he published extensively on various languages, including detailed studies on the Lappish dialect of Finland. He was particularly interested in the “musical accent” of this language. Poirot provided acoustic data from recordings of a diverse range of examples, measured in vibrations per second (Hz). To analyze the melodic movements, he employed a method adapted from Pipping, a specialist in Finno-Ugric languages also based in Helsinki. This method involved calculating pitch intervals as a percentage of the lowest frequency in a given comparison [17, 18]. Notably, Pipping’s method has not gained traction outside of Finland. In his writings, Poirot also reflected on the variety of pitch scales then in use, including tempered scales.

None of these leading phoneticians of the time utilized a logarithmic pitch scale. To investigate whether any others may have, we examined the two primary journals publishing experimental phonetics at the time: *La Revue de Phonétique* (edited by Rousselot & Pernot, 1911–1929) and *VOX* (edited by H. Gutzmann, Panconcelli-Calzia, 1913–1922). Our search did not yield a single article featuring the logarithmic scale. While our research is not exhaustive and we cannot definitively conclude that the logarithmic scale was used by phoneticians at the beginning of the 20th century, it appears to have been rare and not a dominant trend in the field.

While multiple methods for pitch analysis were proposed and tested, none adopted a log-based scale as a standard for visualization. It is within this context of methodological fragmentation and limited systematization that Fu Liu’s work stands out: he appears to be the only phonetician of his generation to have built a complete graphical system grounded in logarithmic pitch mapping. This system was fully quantified and based on the twelve-tone equal temperament, an idea that resonates not only with Western tuning theory but also with the sixteenth-century Chinese mathematician and music theorist, Zaiyu Zhu⁴. Liu’s system was more than a pragmatic solution to the challenge of pitch visualization; it represented a paradigm shift in how tonal variation could be captured, compared, and reproduced. The following section reconstructs this system from its conceptual foundations to its manual implementation, examining how it reconciled

⁴Zaiyu Zhu is the modern pinyin romanization of 朱载堉. He is also referred to as Chu Tsai-yü in some earlier English-language publications.

mathematical rigor, acoustic theory, and pedagogical design long before the digital age.

3 Fu Liu’s Logarithmic Scale Paradigm: Theory, Implementation, and Impact

3.1 From Frequency to Form: Conceptualizing a Logarithmic Pitch Scale

As previously discussed, one of the key challenges for early twentieth-century phoneticians was how to visualize tonal variation in a way that captured both acoustic precision and perceptual relevance. Conventional musical systems such as just intonation provide a foundation for understanding pitch intervals, but these intervals are not equal in terms of physical frequency (e.g. a pure fifth spans 702 cents, while a pure major third spans 386 cents). When plotted on a linear frequency axis, the same interval occupies different distances depending on its register, resulting in perceptual distortion of pitch contours.

A typical example can be found in C. B. Bradley’s 1911 tonal study of Siamese [4], in which both frequency and musical notes were marked on the chart. However, the actual curves were plotted based on frequency values alone, and the musical notes served only as approximate references, failing to account for the non-linear nature of human pitch perception.

Fu Liu recognized that a more systematic approach was needed for tonal languages such as Chinese, where small intervallic differences carry phonological significance. During his training in London and Paris, he adopted a logarithmic axis for pitch plotting, constructing what he called a “fixed scale”. In this system, each frequency ratio corresponded to a fixed graphical distance: an octave (2:1) occupied the same vertical span regardless of its absolute frequency, and each semitone in the twelve-tone equal temperament was equally spaced.

He adopted the concept of equal temperament from Western music theory, in which the octave is divided into twelve equal parts, each semitone corresponding to a frequency ratio of $2^{1/12}$. When plotted on a \log_2 scale, this structure ensures that pitch contours retain their shape across registers. Fu Liu’s adoption of this idea provided a unified coordinate system for tonal comparison, allowing tone curves to be measured, drawn, and compared with precision.

It is worth noting, however, that the idea of dividing the octave into equal intervals had a parallel mathematical formulation in China. In the late sixteenth century, Ming dynasty mathematician and music theorist Zaiyu Zhu systematically derived the structure of the twelve-tone equal temperament in his treatise *Lüxue Xinshuo*, calculating the frequency ratios of semitones with remarkable precision (up to 24 decimal places). Deeply versed in classical Chinese pitch theory and simultaneously trained in modern acoustic phonetics, Fu Liu’s use of a logarithmic frequency axis in tonal visualization can be seen as a conceptual transfer the principles of equal-tempered tuning into the domain of speech acoustics. His system thus fused traditions across continents and centuries to construct a novel method for tonal visualization.

3.2 Designing a Quantitative Pitch Axis: As early as Zaiyu Zhu

To realize his pitch visualization system, Liu constructed a pitch axis with uniform increments designed to reflect equal perceived differences in pitch. Such an idea was not without precedent: it can be traced back to both Chinese and Western traditions.

In China, as previously mentioned, Zaiyu Zhu proposed the twelve-tone equal temperament system in the late sixteenth century, dividing the octave into twelve equal intervals (semitones) and computing their frequency ratios with high precision. This system also resolved a long-standing problem in Chinese tuning theory, known as the “inability to

restore Huangzhong” (黄钟不能还原) in the *sānfēn sǔnyì* (三分损益) method⁵ [16].

In the West, similar attempts were made in the early seventeenth century by Italian musician Vincenzo Galilei and Flemish mathematician Simon Stevin [24], but their proposals were not widely accepted [2]. It was not until the eighteenth century, driven by the needs of keyboard modulation, that twelve-tone equal temperament became the dominant tuning method [6, 25]. Later, in 1885, British scholar Alexander J. Ellis introduced the concept of “cents” in his work *On the Musical Scales of Various Nations*, defining the octave as 1200 cents and each semitone as 100 cents [5], thereby enabling fine-grained numerical representation of pitch.

Fu Liu applied the principles of equal temperament and the cent system in constructing his pitch axis. He subdivided each semitone into five parts, such that each small step represented 20 cents, with a frequency ratio of approximately $2^{1/60} \approx 1.0116$. He selected the French standard reference pitch of $la_3 = 435$ Hz (corresponding in pitch to A_4 in the Anglo-German system, though slightly lower in frequency) as the baseline frequency and calculated downward to $ut_0 (C_1) = 32.3$ Hz, thus generating a frequency table (figure 6) covering approximately four octaves.

In his doctoral dissertation [9], Liu compiled Table I, which lists the frequencies (in Hz) and common logarithms (\log_{10})⁶ for each semitone and their respective sharps and flats (e.g., *ré#*). He noted that since the frequency ratio between adjacent semitones is constant at $2^{1/12}$, their logarithmic difference should also be constant. For example, the \log_{10} difference between 435 Hz (la_3) and the next semitone at 410.6 Hz is approximately 0.0251. This constant was denoted as D . Dividing D by 5 yields the difference between adjacent 1/5-semitone intervals ($d = 0.00502$), corresponding to 20-cent steps.

To facilitate pitch plotting, Liu extended Table I by calculating the frequencies for each 1/5-semitone increment using the constant ratio $r = 2^{1/60}$. Starting from $ut_0 = 32.3$ Hz, he generated a complete set of values such as 32.7 Hz, 33.1 Hz, and so on, up to $la_3 = 435$ Hz (see Table II). These values established a direct mapping between acoustic frequency and vertical position on graph paper, where each 1 mm step represented a 20-cent increment. For example, a measured f_0 of 125 Hz could be located between $ut_2 (C_3)$ and $si_1 (B_2)$ at approximately 40–60 cents, enabling rapid and precise plotting without the need for real-time logarithmic calculation.

3.3 Tonal Contour Drawing Procedures and Methodological Implications

Once the tables were complete, Fu Liu established a standard procedure for drawing tonal curves. In processing tonal data, he used a kymograph to record vibration trajectories, then measured the frequency for each cycle, matched it to the corresponding value on his logarithmic scale, and marked the points along the time axis, as shown in figure 7. These points were then smoothly connected manually to form a continuous pitch contour.

Liu emphasized that the smoothing process was crucial: the curve should not be mechanically interpolated point-by-point, but drawn according to the overall trend, thereby

⁵*Huangzhong*, literally “Yellow Bell,” is the foundational pitch in traditional Chinese music theory, serving as the reference pitch from which the entire tuning system is constructed. Under the *sanfen sunyi* method, which constructs pitches by iteratively multiplying or dividing by 3/2 or 4/3, the system fails to return to the original *Huangzhong* pitch after a full cycle of intervals. Zaiyu Zhu’s equal temperament reform addressed this issue by proposing a mathematically closed system where each semitone is spaced equally on a logarithmic scale.

⁶Liu used base-10 (\log_{10}) logarithms to tabulate frequency values, as shown in figure 6. While this differs from the \log_2 transformation more commonly used in modern phonetics, the intervallic structure remains identical due to the constant frequency ratio per step. For consistency with modern practice, we adopt \log_2 notation elsewhere in this paper, while retaining Liu’s original \log_{10} values where relevant.

TABLE I					TABLE I (suite)				
DT.	VD.	L.	D.	d.	DT.	VD.	L.	D.	d.
la ₁	433,0	2,6385			si ₁	122,1	2,0867		
sol ₁	410,6	2,6134	0,0251	0,00502	la ₂	115,2	2,0613	0,0252	0,00504
sol ₂	387,5	2,5883	0,0251	0,00502	la ₃	108,8	2,0367	0,0248	0,00495
fa ₂	365,8	2,5633	0,0250	0,00500	sol ₃	102,7	2,0113	0,0252	0,00504
fa ₁	345,3	2,5382	0,0251	0,00502	sol ₂	96,9	1,9863	0,0249	0,00498
mi ₂	325,9	2,5131	0,0251	0,00502	fa ₃	91,5	1,9614	0,0254	0,00508
ré ₂	307,6	2,4880	0,0252	0,00504	fa ₂	86,3	1,9360	0,0248	0,00496
ré ₁	290,3	2,4628	0,0250	0,00500	mi ₁	81,5	1,9112	0,0253	0,00506
ut ₂	274,0	2,4378	0,0250	0,00500	ré ₂	76,9	1,8859	0,0250	0,00500
ut ₁	258,7	2,4128	0,0252	0,00504	ré ₁	72,6	1,8609	0,0252	0,00504
si ₂	244,1	2,3876	0,0252	0,00504	ut ₂	68,5	1,8357	0,0248	0,00496
la ₃	230,4	2,3624	0,0249	0,00498	ut ₁	64,7	1,8109	0,0256	0,00512
la ₂	217,5	2,3375	0,0251	0,00502	si ₂	61,0	1,7853	0,0249	0,00498
sol ₃	205,3	2,3124	0,0250	0,00500	la ₂	57,6	1,7604	0,0248	0,00496
sol ₂	193,8	2,2874	0,0252	0,00504	la ₁	54,4	1,7356	0,0253	0,00510
fa ₃	182,9	2,2622	0,0252	0,00504	sol ₂	51,3	1,7101	0,0253	0,00506
fa ₂	172,6	2,2370	0,0251	0,00502	sol ₁	48,4	1,6848	0,0249	0,00498
mi ₁	162,9	2,2119	0,0250	0,00500	fa ₂	45,7	1,6599	0,0241	0,00488
ré ₂	153,8	2,1869	0,0249	0,00498	fa ₁	43,2	1,6353	0,0259	0,00518
ré ₁	145,2	2,1620	0,0253	0,00506	mi ₂	40,7	1,6096	0,0241	0,00482
ut ₂	137,0	2,1367	0,0251	0,00502	ré ₂	38,5	1,5853	0,0256	0,00512
ut ₁	129,3	2,1116	0,0249	0,00498	ré ₁	36,3	1,5599	0,0246	0,00492
					ut ₂	34,3	1,5353	0,0261	0,00522
					ut ₁	32,3	1,5092		

Figure 6. Fu Liu's Table I: Frequencies and logarithmic values of equal-tempered semitones, including base-10 logarithmic differences (D) and their subdivisions (d). *Liu 1925*, p. 14-15.

filtering out noise or error introduced by equipment limitations. This method was applied in his 1925 doctoral dissertation *Étude expérimentale sur les tons du chinois* [9], where he presented tone curves for ten Chinese dialects, including Beijing, Jiangyin, and Cantonese, across monosyllabic, disyllabic, and continuous speech contexts. The results demonstrated a remarkable level of precision and comparability, allowing typological and stylistic analyses to be conducted on a shared quantitative basis.

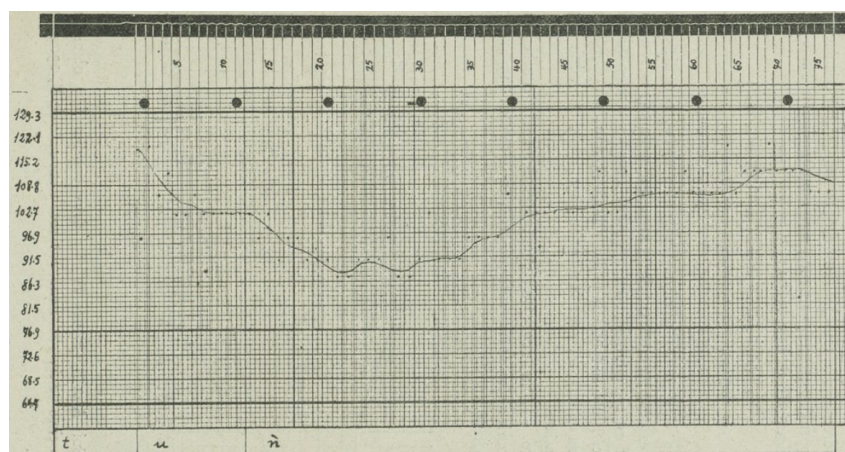


Figure 7. Smoothed pitch contour manually drawn using Fu Liu's logarithmic pitch scale on grid paper *Liu 1925*, p. 19.

Liu's entire workflow constituted a reproducible methodological framework. By standardizing the grid paper (with 20-cent vertical divisions) and using a consistent lookup table, researchers could plot and compare tonal data on a common coordinate system. Though developed decades before the advent of digital tools, it captured key elements of what would later become standard in phonetic analysis: perceptual scaling, graphic precision, and inter-speaker comparability.

To be clear, Liu’s method did not directly influence the development of software such as Praat [3], nor was it informed by the psychoacoustic pitch scales (such as the Mel or Bark scales) introduced later in the twentieth century. However, the mathematical logic of his scale, which expresses pitch in semitone steps relative to a fixed reference frequency (435 Hz), is formally equivalent to the logarithmic conversions widely used in modern systems. In this sense, Liu’s method can be seen as an independent and remarkably early solution to the problem of how to represent tone variation in both acoustic and perceptual terms. Rather than anticipating specific tools, his work prefigured a broader methodological shift in phonetics: from descriptive approximations to precise, repeatable measurements. By building a full-scale, log-based coordinate system for pitch drawing by hand, he laid empirical foundations for later cross-linguistic tonal comparisons.

Liu’s contributions extended beyond experimental techniques. In works such as *Pípá jí tā zhǒng xián yuèqì zhī děnglǜ dìng pǐn fǎ* [10], *Shí’èr děnglǜ de fāmíng zhě Zàiyǔ Zhū* [13], and *Cóng wǔyīn liùlǜ shuō dào sānbǎi liùshí lǜ* [11], he traced the evolution of tuning systems in Chinese traditions, showing deep awareness of both mathematical and cultural dimensions of pitch. His work exemplifies an early and original attempt to bridge acoustic phonetics, perceptual modeling, and cross-cultural theoretical insight.

4 Conclusion

This study has reconstructed the theoretical rationale and technical implementation of the logarithmic framework developed by Fu Liu for tonal analysis. Situated within the broader context of early 20th-century phonetics, Liu’s approach is distinguished by its integration of a logarithmic scale and twelve-tone equal temperament, a combination largely absent from the major European phonetic traditions of the time. Rather than constituting a derivative or marginal contribution, Liu’s system offers an original and integrative attempt to formalize pitch measurement with internal consistency and reproducibility. Bridging musical acoustics, phonetic theory, and empirical observation, it embodies methodological principles that would only become widespread in phonetics decades later. As such, Liu’s legacy not only enriches our historical understanding of phonetic science, but also calls for a more inclusive reassessment of contributions that emerge outside the mainstream European canon.

5 Acknowledgments

Jingyi Sun is currently funded by the China Scholarship Council (Grant No. 202208410095), with additional support from the CNRS (Centre National de la Recherche Scientifique).

References

- [1] ASHBY, M. The first spoken intonation corpus (1909): A re-assessment of Daniel Jones’s Intonation Curves. In *10th International Conference on Speech Prosody* (2020), pp. 121–125.
- [2] ASSELIN, P.-Y. *Musique et tempérament*, vol. 1. Éditions Costallat, 1984.
- [3] BOERSMA, P., AND VAN HEUVEN, V. Speak and unSpeak with PRAAT. *Glott International* 5, 9/10 (2001), 341–347.
- [4] BRADLEY, C. B. Graphic analysis of the tone-accent of the Siamese language. *Journal of the American oriental society* 31, 3 (1911), 282–289.
- [5] ELLIS, A. J. *On the musical scales of various nations*. Journal of the Society of arts, 1885.
- [6] GRENFELL, M. T., AND DANBURY, C. The Development of The Equal Temperament Scale.

- [7] JONES, D. *Intonation curves: A collection of phonetic texts, in which intonation is marked throughout by means of curved lines on a musical staff*. BG Teubner, 1909.
- [8] JONES, D. *An outline of English phonetics*. BG Teubner, 1922.
- [9] LIU, F. *Étude expérimentale sur les tons du chinois*. Paris: Société d'édition les Belles-Lettres; Pékin: Presses de l'Université nationale de Pékin, 1925.
- [10] LIU, F. 琵琶及他种弦乐器之“等律”定“品”法 (On Equal-Tempered Fret Placement in the Pipa and Other String Instruments). *Weekly Journal of the Institute for National Studies, Peking University* 2, 17 (1926).
- [11] LIU, F. 从五音六律说到三百六十律 (From the Five Yin and Six Lü to the 360-Lü System). Public lecture delivered at the Institute for National Studies, Peking University, 1927.
- [12] LIU, F. 声调之判断及“声调判断尺”之制造与用法 (On the Judgment of Tones and the Construction and Use of the Tone Inference Ruler). *Bulletin of the Institute of History and Philology, Academia Sinica* 7 (1928–1930).
- [13] LIU, F. 十二等律的发明者朱载堉 (Zhu Zaiyu, the Inventor of the Twelve-Tone Equal Temperament). *Institute of History and Philology, Academia Sinica*, 1932, pp. 279–310.
- [14] MÜNNING, M. The ‘French Connection’—Chinese Linguists in Paris.
- [15] PELLIN, T. The sweet revolutionaries: The Chinese revolution in grammar studies and Henry Sweet. *Language & History* 54, 1 (2011), 35–57.
- [16] PENG, B. On the Western reception of Prince Zhu Zaiyu’s music theory from the eighteenth to the twentieth century. *International Communication of Chinese Culture* 10, 2 (2023), 133–149.
- [17] POIROT, J. Recherches expérimentales sur le dialecte lapon d’Inari. Accent musical. *Finnisch-Ugrische Forschungen* 1–3 (1904), 11–57.
- [18] POIROT, J. Recherches expérimentales sur le dialecte lapon d’Inari. *Finnisch-Ugrische Forschungen* (1905), 153–230.
- [19] ROUSSELOT, P.-J. *Les modifications phonétiques du langage étudiées dans le patois d’une famille de Cellesfrouin (Charente)*. 1891.
- [20] ROUSSELOT, P.-J. *Principes de phonétique expérimentale*. H. Welter, 1908.
- [21] ROUSSELOT-L’ABBÉ, P. *Principes de phonétique expérimentale. Tom II*, Welter, éditeur (1901).
- [22] SCRIPTURE, E. W. *The elements of experimental phonetics*, vol. 22. C. Scribner’s Sons, 1902.
- [23] SCRIPTURE, E. W. *The study of English speech by new methods of phonetic investigation*, vol. 11. H. Milford, Oxford University Press, 1923.
- [24] STEVIN, S. Simon Stevin, 1970.
- [25] STURM, F. A Clear and Practical Introduction to Temperament History, Part 1: Introduction. *Piano Technicians Journal* 53, 5 (May 2010), 13–19. Published by the Piano Technicians Guild.
- [26] SUN, J., GAO, J., AND RIALLAND, A. Pioneering works in modern Chinese phonetics: Fu Liu’s tonal studies and other Chinese scholars at the University of Paris (1920s-1940s). In *Proc. HSCR 2024* (2024), pp. 56–65.

PRESERVING X-RAY FILMS OF SPEECH PRODUCTION

K.G. Munhall¹, F. DiBona², D. Norman¹, J. Szutka²

¹ *Dept. of Psychology, Queen's University, Kingston, Canada,*

² *Vulnerable Media Lab, Queen's University, Kingston, Canada*

munhallk@queensu.ca

Abstract: X-ray films played an important role in speech research in the last century. The images of the moving vocal tract are unsurpassed for the information they provide about speech coordination and phonetic categories. Unfortunately, these films are in danger of loss through neglect and deterioration of the films themselves. The present paper gives a summary of our attempt to preserve a large set of these films.

1 Introduction

Measuring the vocal tract during speech has been one of the fundamental challenges of the phonetic sciences. Speech sounds are produced internally by the changing shape of the oral cavity and thus are only indirectly visible from the face and throat. The complex morphology of the tongue and oropharyngeal region and the rapidity of speech motion adds to the problem of inaccessibility. A variety of ingenious gadgets and transducers [1] were developed over the years to characterize different aspects of talking but the introduction of radiographic images of speech was a breakthrough. The static images that were initially recorded showed much about vowel and consonant articulation (e.g., [2]) during sustained postures. However, the use of X-ray movies of fluent speech beginning in the 1950s revealed speech coordination in all its remarkable complexity.

From the initial uses of X-rays in speech research, there was awareness of the potential health risks. Russell [2] commented on experiencing radiation burns. By the mid 1970s most X-ray studies of speech in normal subjects stopped. There was a consensus that the radiation dosage received by participants in X-ray research studies was not justifiable. The films recorded before this time were set aside. Some were stored in labs and were rarely available to researchers and students. Others were misplaced or destroyed. These films still display the best view of speech articulation and in the 1990s, we decided to preserve and make available a group of such films [3].

Twenty-five films were transferred to LaserDisc and distributed upon request without charge. Twenty-four of these films were from a large project on the study of French and English phonetics led by Professor Claude Rochette at Université Laval, Québec (e.g. [4]). The remaining film was recorded in Stockholm, Sweden by Drs. S. Öhman and K. Stevens [5]. These LaserDiscs were widely distributed for use in such diverse areas as research, teaching, art installations, music videos, animation, and so on.

This paper describes a continuation of our preservation project. We have digitized a large set of films using a high-resolution format. Our choice of LaserDisc in the 1990s seemed like a good idea at the time but clearly, we selected a technology with no future. In addition to the change in format our work now has a new urgency. Existing X-ray films are degrading and may soon be lost to a breakdown of the films themselves.

Prior to about 1951, all movie film stock had a nitrate base. These cellulose nitrate films were extremely flammable and many films from the early days of commercial movies have been lost to fires. As the use of X-ray movies was beginning in the early 1950s, film stock

changed to one with an acetate base. These films were much more stable and not a fire risk. However, these acetate films will also deteriorate depending on the heat and moisture conditions of storage. In hot and humid conditions acetic acid or vinegar is released. This degradation process is called vinegar syndrome. The physical result on the films includes irregular shrinkage which produces warping in the width and length of sections of the film. As the degradation advances, the films can lose flexibility, cracks in the emulsion occur and breaking and flaking of the film surface occurs [6].

2 The Films

Currently we have collected a set of 58 X-ray movies (see Table 1). The provenance of some of the films is vague and we cannot say with certainty which publications some of these films were reported in. In this set of films, some already show signs of advanced vinegar syndrome which reinforces our decision to move forward on this project now. We reached out to others in the field to increase this film set but were disappointed to learn that many notable films were missing. The Iowa films that were the basis of studies by Moll, Kent, Daniloff and others were destroyed. A large number of the Université Laval films were destroyed because of vinegar syndrome. Gunar Fant's films which informed his modelling for the acoustic theory of speech production are currently missing with no trace.

Table 1. Inventory of digitized films

Number of films	Source/Researcher	Language	Date recorded
5	Delattre	3 films of English, 1 film of English with a small section of French, Spanish, and German, 1 film of Spanish (missing audiotape)	1966
4	Haskins	3 films of English and 1 film of Arabic	1961, 1962
1	Houde	English	1966
3	KNS Film Öhman & Stevens	English	1962
44	Laval	15 films of English, 24 films of French, 5 films currently without audiotape	1974
1	Truby	English	1958

The films labelled Delattre in Table 1 were recorded in the Speech Synthesis Project led by Pierre Delattre at the University of California, Santa Barbara. The films appear to be part of the study of the production of /r/ and other consonants in different languages and dialects (e.g., [7]). The Haskins Laboratories films were part of a multi-language series of recordings (Mandarin, Syrian Arabic, Hungarian and Russian) directed by Arthur Abramson and Franklin

Cooper. The films were recorded at Columbia-Presbyterian Hospital in New York City and shown in slow motion with a stretched speech soundtrack [8]. The Houde film was recorded by Robert Houde as part of his doctoral thesis [9]. The so-called KNS film was recorded at the Wenner-Gren Research Laboratory at Nortull's Hospital, Stockholm, Sweden. We have three copies of the film including the version analyzed by Joseph Perkell in his landmark monograph [10]. We copied this film in the 1990s but still included it in this project to preserve a high-resolution digital version. The Laval films preserved here are different ones than those archived to LaserDisc in the 1990s. We hoped to include all the Laval films in this project but unfortunately many had been destroyed. Overall, the Université Laval project is a remarkably large collection of films in French and English. The Truby film is a scientific/educational film created in Stockholm in 1958 using some of Truby's early X-ray studies. The film, Correlation of Cineradiographic and Visual Acoustic Analyses of Speech, was presented at the Fourth International Congress of Phonetic Sciences [11].

3 Film Preservation

There are two parts to the preservation of the films. The first is to establish storage conditions that safeguard the original films. Queen's University Archives has climate-controlled storage and have accepted our film collection. The films were assessed for vinegar syndrome and the ones showing moderate to advanced signs of this degradation were frozen in the cold storage facility at the Archives. The testing for vinegar syndrome can be done crudely by opening the film canister and smelling for any indication of vinegar. A more objective approach is to use A-D strips produced by the Image Permanence Institute at the Rochester Institute of Technology. We used these specially coated paper strips that are designed to detect acetate deterioration and assess its extent.

The second part of the preservation process involved the creation of digital copies of the films. The digitizing of the films and acoustic recordings was carried out at the Vulnerable Media Lab at Queen's University using a Lasergraphics ScanStation. Archival quality TIFF images (5120X3840) were made for each film frame. Uncompressed .WAV files were recorded to preserve the audio signals. Since the file directories of these high-resolution images are quite large, smaller access files (.MP4 movies using H.264 video compression format) were produced to allow for ease of viewing and file sharing. The high-resolution images were concatenated into movies and synchronized with the audio files playing at recording (natural) speed using Adobe Premiere Pro. In the .MP4 movies, the images were modified in two ways. The images were compressed using the H.264 codec and resolution was decreased to 1920X1080 and for movies that were filmed at highspeed, frames were decimated to allow a standard playback of 24 frames per second and synchronization with the audio track.

Synchronization is a challenge for some of the films. Several of the films such as KNS, the Laval films and the Houde film have separate audio tapes and films. The audio tapes can become stretched from use over the years. For some films there is a timing signal on both the film and audiotape, but it is present only at a single time point. Some of the films are 16 mm films with optical or magnetic soundtracks on the film and there is no issue with synchronization with these films. The .MP4 movies should be viewed as being only loosely synchronized for the purpose of viewing. Research use of the films to study the vocal tract motions and accompanying acoustics will require researchers to establish their own synchronous signals.

4 Conclusion

Our presentation will be a brief overview of the steps taken to preserve a set of 58 X-ray films of speech production. Examples of the films will be shown. However, the main purpose of our presentation will be to draw attention to this vulnerable resource from the history of research into articulation. The films recorded in the 1950s-1970s are a unique set of speech images that will never be recorded again.

5 References

- [1] HARDCASTLE, W. J.: Instrumental investigations of lingual activity during speech: a survey, *Phonetica*, 29(3), pp. 129-157, 1974.
- [2] RUSSELL, G.O.: *The Vowel: Its Physiological Mechanism as Shown by the X-Ray*. Columbus: The Ohio State University Press, 1928.
- [3] MUNHALL, K. G., E. VATIKIOTIS-BATESON, and Y. TOHKURA: X-ray film database for speech research, *The Journal of the Acoustical Society of America*, 98(2), pp. 1222-1224, 1995.
- [4] ROCHETTE, C.E.: *Les Groupes de consonnes en français: étude de l'enchaînement articulatoire à l'aide de la radiocinématographie et de l'oscillographie*. Les presses de l'université Laval, 1973.
- [5] ÖHMAN, S. E., and K.N. STEVENS: Cineradiographic studies of speech: procedures and objectives, *The Journal of the Acoustical Society of America*, 35(11_Supplement), pp. 1889-1889, 1963.
- [6] NATIONAL FILM PRESERVATION FOUNDATION: *The Film Preservation Guide: The Basics for Archives, Libraries, and Museums*. San Francisco: National Film Preservation Foundation, 2004.
- [7] DELATTRE, P., and D.C. FREEMAN: A dialect study of American r's by x-ray motion picture. *Linguistics*, 44, pp. 29-68, 1968.
- [8] ABRAMSON, A. S., and F.S. COOPER: Slow-Motion X-Ray Pictures with Stretched Speech as a Research Tool, *The Journal of the Acoustical Society of America*, 35(11_Supplement), pp. 1888-1889, 1963.
- [9] HOUDE, R. A.: *A study of tongue body motion during selected speech sounds*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI, 1967.
- [10] PERKELL, J.S. *Physiology of speech production: results and implications of a quantitative cineradiographic study*. MIT research monograph, 53. Cambridge, Mass.: MIT Press, 1969.
- [11] TRUBY, H. M.: Synchronized cineradiography and visual-acoustic analysis. In *Proceedings of the Fourth International Congress of Phonetic Sciences*, Mouton and Co., The Hague, pp. 265-279, 1962.

6 Acknowledgements

The work was funded by a grant to the first author by the Natural Sciences and Engineering Research Council of Canada. Keanna Mackinnon did the initial sorting of our collection of films and audio tapes. We thank Jeremy Heil at the Queen's University Archives who has been instrumental in providing a permanent home for the original films and making the digital copies available to future researchers.

ULTRASOUND IN HUNGARIAN SPEECH RESEARCH: METHODS AND APPLICATIONS

Alexandra Markó¹, Andrea Deme^{2,1,3}, Kornélia Juhász^{3,1,2}, Tekla Etelka Grácz^{3,1}

¹ *MTA–HUN–REN NYTK Lendület “Momentum” Neurophonetics Research Group, Hungary,*

² *ELTE Eötvös Loránd University, Dept. of Applied Linguistics and Phonetics, Hungary,*

³ *HUN–REN Hungarian Research Centre for Linguistics, Hungary*

marko.alexandra.phd@gmail.com

Abstract: This paper presents the efforts of Tamás Gábor Csapó (1985–2024), who – due to his extensive international connections – played a key role in the initialization of ultrasound tongue imaging in Hungary as well as in the development of Silent Speech Interface (SSI) systems and other speech technology applications that support the lives of people with speech disabilities.

1 Introduction

This paper is dedicated to the memory of our talented colleague and unforgettable friend, Tamás Gábor Csapó (1985–2024), without whom we could not speak about ultrasound tongue imaging studies of Hungarian speech. Thanks to his extensive international connections, Tamás Gábor Csapó mastered the Ultrasound Tongue Imaging (UTI) technique, and played a key role making the first ultrasound device available in Hungary. He worked extensively to ensure that the most pioneering methodological research could be conducted in international joint works. He was always mindful of helping disadvantaged people, and thus he immediately became involved in the development of Silent Speech Interface (SSI) systems and other speech technology applications that facilitate the lives of people with speech disabilities. His work in the field of ultrasound analysis and applications is widely acknowledged both in Hungary and internationally. It is a huge loss for us and our discipline that he was given such a short time.

The paper is organized as follows: in session 2 we give a short overview of main characteristics of UTI and in session 3 about the historical background of dynamic articulatory studies of Hungarian speech. We present some methodological issues from the scope of Csapó’s contributions to the field in session 4, and finally, some UTI-based applications will be mentioned worked on by Csapó in session 5.

2 Ultrasound in articulatory analysis

UTI serves for the examination of the surface of the tongue (however, under certain conditions the palate can also be drawn with its help, and occasionally it is also used to track the position of the larynx [28]). Ultrasound is an imaging technique that emits ultra-high frequency sound waves generated by exciting a piezoelectric crystal. The sound waves pass through the tissue and reflect off its surface [41]. The ultrasonic transducer placed under the chin creates an image of the tongue. The typical result of (two-dimensional) ultrasound recordings is a series of grayscale images on which the tongue’s surface contour appears with greater brightness than the surrounding tissues and air.

The application of ultrasound technique for examining the function of speech organs was suggested in the late 1960s by Kelsey, Minifie, and Hixon [27], primarily as an alternative to examinations performed with X-rays, which were considered unjustifiably risky for healthy people due to radiation exposure. By contrast, ultrasound – as the authors stated (p. 564) – is

safe, fast, and does not cause discomfort to the subject. Accordingly, ultrasound has increasingly been used in phonetic research, especially in the last twenty years [28]. On the international popularity list of articulatory research methods, UTI occupies the second place after Electromagnetic Articulography (EMA) [28].

The primary advantage of ultrasound is its ability to capture the movement of the entire tongue surface, as opposed to only a few specific points of the tongue, as is the case with EMA. It is also easy to use, affordable, and allows for recording with high resolution (up to 800×600 pixels) and high speed (up to 100 frames/s) [10]. Additionally, it is well applicable to the vast majority of speakers because it does not cause significant discomfort to participants. Further advantages include its portability owing to its compact size and lightweight design. Thus, speakers from vulnerable populations (e.g., children) do not need to be examined in a phonetic laboratory; experiments can be recorded in their natural environment. In addition, UTI can also be easily combined with other instruments, primarily with an electroglottograph (EGG). This allows for simultaneous recording and examination of tongue articulatory activity and glottal function.

By contrast, the UTI's disadvantage is that it only makes the movement of a single speech organ, the tongue, observable. It is quite sensitive to the condition of the speaker's tissues (e.g., fatness, hydration), and to some extent also to body size because it provides a less clear image in larger cavities than in smaller ones (e.g., [10; 41]). Further difficulty is presented by fixing the ultrasound transducer to the moving articulator (jaw), as well as data extraction (documentation and analysis of tongue contours). During articulation, the palate is not visible in the ultrasound image because the ultrasound beam reflects off the air in the oral cavity. However, regarding tongue position, the distance between the tongue and the palate is inevitably important information. During swallowing, the tongue is almost completely in contact with the palate, making the palate contour observable [18]. The surface line of the palate can thus be recorded in the ultrasound image and used as a reference in examining tongue position. Further reference can be provided by the recording of the bite plane [40] which can be recorded by placing a rigid plate between the two rows of teeth, on which the subject bites. By resting the tongue against the bottom of the plate, the tongue contour traces out the bite plane.

Two-dimensional ultrasound (which has been used so far for examining Hungarian speech) records the movement of the tongue's sagittal midline [42]. A helmet is used to position the ultrasound transducer under the chin (see an example in the left image of Figure 1). Simultaneously with the ultrasound recording, speech is recorded with a microphone clipped to the helmet. Synchronized recording of video and audio, the display of recordings, and the analysis are made possible or supported by some software. The right panel of Figure 1 shows an example of this: in certain windows of the Articulate Assistant Advanced (AAA) software, the ultrasound image, speech waveform, FFT spectrum, and spectrogram are visible simultaneously [3]. In addition, as shown at the bottom of the panel, the AAA software is also suitable for automatic tongue contour tracking, thus capable of automatically scanning the outline of the tongue's surface on the ultrasound images (with some errors).

Although the ultrasound makes the tongue surface contour visible (Figure 1, right panel), this contour cannot be directly analyzed. To be able to examine the tongue shape and its changes, the tongue contour (or its data) must be extracted from the recorded series of images using some method with the help of a dedicated software. This extraction can be performed manually: during this process, the tongue contour is drawn onto the ultrasound image displayed on the screen by moving the mouse. However, this is extremely time-consuming. The other possibility is to attempt to identify the tongue contour using automatic methods (see the red lines in the bottom image series at right panel of Figure 1). However, the accuracy of these methods does not yet reach that of manual contour extraction. Therefore, human checking and correction are absolutely necessary when automatic methods are used [48].

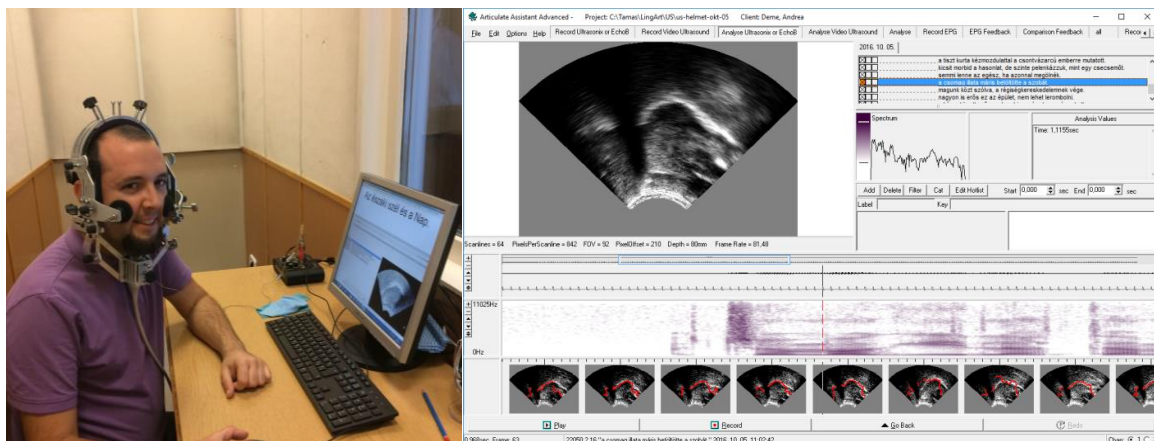


Figure 1. Tamás Gábor Csapó in the helmet [2] fixing the ultrasound transducer (left), and the screen capture of the Articulate Assistant Advanced (AAA) software (right, [3; 10 p. 343; 13 p. 3735]).

3 Analysis of Hungarian speech with UTI: Historical background

Regarding the Hungarian language, prior to the domestic appearance of UTI, only a few studies on lingual articulation were based on dynamic data (i.e., not just static images) and recorded the movement of speech organs in real time (instead of documenting the imprint afterwards, such as photopalatography). János Lotz in 1966 [30], Tamás Szende in 1974 [43], and Kálmán Bolla in 1981 [5; 6] investigated tongue gestures using X-ray film technique. In the early 2000s, various studies were conducted using EMA in foreign laboratories [31; 16; 17]. It is worth noting that despite the fact that the palatography technique has a nearly 140-year tradition in Hungary, and many different studies were conducted with the static (photopalatography) methodology (see [20; 33]), electropalatography, which is suitable for the analysis of the dynamic aspect of tongue gestures, has not been used in our country so far.

Although the application of UTI in articulatory research is documented in the international literature since the 1960s [27], the first UTI study on Hungarian speech was published in 2007 by Stefan Beňuš and Adamantios I. Gafos [4]. Their experiment was recorded at Haskins Laboratories using EMA (3 speakers) and ultrasound (1 speaker). The researchers studied a specific phenomenon of the Hungarian language, vowel harmony, and within that, the articulatory characteristics of transparent vowels.

Tamás Gábor Csapó mastered the UTI technique during a Fulbright Scholarship at Indiana University, Bloomington, under the supervision of Steven Lulich, in 2014.

In domestic research, UTI studies began in 2016, when due to a grant support the MTA–ELTE Lendület Lingual Articulation Research Group was able to afford higher-value articulatory instruments. As a founding member of the research group, Tamás Gábor Csapó played a significant role in the domestic implementation of the UTI technique. The research group’s studies focused on phenomena of connected Hungarian speech, such as the articulatory differences between geminate and singleton stops and fricatives regarding tongue position [35; 36; 37]; the longitudinal change in the articulatory pattern of voiceless sibilants in children [23]; and the differences in the production characteristics of voiced and voiceless obstruents [22; 24; 25].

Tamás Gábor Csapó has conducted pioneering work—even in an international context—in the development of UTI data analysis methodology, silent speech interfaces, and other speech technology-based digital applications (e.g., [15]). Several individual and collaborative projects led by him further advanced these areas: 2017–2021: *Silent Speech Interface based on articulatory movements* (NKFIH FK-17 research project); PI: Tamás Gábor Csapó (Budapest University of Technology and Economics); 2018–2021: *Development of silent speech interfaces and investigation of speaker dependency using ultrasound tongue imaging and electromagnetic articulography* (NKFIH PD-18 research project); PI: Tamás Gábor Csapó (Budapest University

of Technology and Economics); 2022–2026: *Analysis of articulation and brain signals for speech-based brain-computer interfaces* (NKFIH FK-22 research project); PIs: Tamás Gábor Csapó, Dávid Sztahó (Budapest University of Technology and Economics).

Drawing on the ultrasound recordings collected by Tamás Gábor Csapó at the MTA–ELTE Lendület Lingual Articulation Research Group laboratory, researchers from the University of Miskolc and the University of Debrecen also conducted measurements and comparative analyses aimed at harmonizing and integrating MRI and UTI data (see e.g., [45; 46; 47]).

4 Methodological issues

4.1 Tongue contour tracking

Two-dimensional ultrasound images are grayscale image sequences in which the tongue surface contour appears as a line of greater brightness compared to its surroundings. However, this line alone is not suitable for quantitative analysis. In order to perform analyses on the tongue contour (for example, comparing tongue shapes, taking measurements regarding tongue curvature), the tongue contour (more precisely, its numerical data, i.e., the coordinates of individual points) must be extracted using a dedicated software. As mentioned, manual contour drawing, although accurate, is extremely time-consuming, while automatic contour extraction is fast, but distorted by errors. The comparison and error analysis of manual contour drawing and various automatic methodologies, and the consequent development of machine techniques, has been being in the focus of the studies from the beginning. One of Csapó's first studies concerning UTI technique, conducted with Steven Lulich, was also targeted at this subject [12]. The aim of the study was to characterize the errors occurring during the determination of midsagittal (mid-sectional) tongue contours from two-dimensional ultrasound image sequences. In the experiment, seven individuals – two experts and five university students – manually drew the contours on a total of 1145 tongue ultrasound images from the speech of four speakers, and the researchers quantified the variability associated with the contours. From these same recordings, tongue contours were also extracted automatically using the EdgeTrak [29], TongueTrack [44], and AutoTrace [19] algorithms, and these were then compared to the manually drawn tongue contours using a quantitative method. According to the results, the quality of the ultrasound image had an effect on the accuracy of manual contour drawing to a greater extent than the skill or expertise of the person drawing the contour. Furthermore, the manually tracked contours were more accurate than that of the machine, and among the algorithms, AutoTrace 3.5 proved to be the most accurate. The authors identified three basic types of errors: 1. the program was unable to track the contour of the fast-moving tongue; 2. errors arising from noise, which could be attributed to the poor quality of the image; and 3. when the program did not track certain parts of the tongue. These error types indicated areas that require further improvement. Further tests were also conducted to identify the error rates (e.g., [9]).

4.2 Processing of tongue contour data

The distance between tongue contours can be quantified using the Nearest Neighbour Distance (NND) method [50] or the LOC_{a-i} index [49], among others. Another type of methodology suitable for comparing tongue contours was developed by Tamás Gábor Csapó and Maida Percival (University of Toronto) for the comparison of Hungarian singleton and geminate consonants [35; 36; 37]. In this method, the tongue contours were rotated into a coordinate system defined by the speaker's bite plane. The contours were then divided into coronal, velar, and pharyngeal regions (based on [39]). For each region, polar coordinates were analyzed with linear mixed regression models to determine in which region the radial distance between the ultrasound probe and the tongue surface differed between the consonants. This radial distance difference was interpreted as reflecting differences in tongue shape and position.

4.3 Variance of data arising from differences in speakers and sessions

As we have seen, ultrasound recordings are variable for numerous reasons, and only one of these is the variability inherently characteristic of articulation. It also follows from the UTI methodology itself that the position of the ultrasound probe differs between two recording sessions, even when using a fixing helmet, but can also change during a single recording for the same speaker. A further variability factor arising from ultrasound is the quality of the recordings and thus the quality (e.g., length) of the tongue contours. Inter-speaker variability is also a commonplace in phonetics. All these factors can compromise the validity of the metrics used for quantifying tongue contours.

To date, no method is available that would solve these problems, i.e., would allow for normalization, alignment, or adaptation of UTI recordings taken at different times and/or with different speakers. Tamás Gábor Csapó worked with Kele Xu on methods that could help solve these problems. In their study published in 2020 [13], they first and foremost attempted to quantify errors arising from differences in ultrasound probe position, and in this context, they saw recording the bite plane as a solution. In their view, based on the bite plane, a reference coordinate system should be established [40] to which the tongue contours can be aligned (rotated) at the appropriate angle, thus becoming comparable. The proposed method involves using Mean Squared Error (MSE) of mean ultrasound images as a metric to identify potential probe movement. The advantage of the method is that it is computationally relatively light and does not require any extra equipment or special setup in the data recording. Moreover, the results of MSE analysis are easy to interpret, and they provide a holistic view to the data from a given recording session at a glance [34].

Further analyses were also conducted on this topic [14], and Csapó planned to develop an automatic classifier that would warn if the position of the ultrasound probe differs within a recording during tongue contour analysis. This classifier would have also assigned confidence intervals to the reliability of the data. He thought that recording the tongue in a neutral position (at the beginning or end of individual utterances) and registering changes in this reference image would also be helpful for the automatic detection of problems arising during the recording.

5 UTI in speech synthesis applications

Speech synthesis traditionally means text-to-speech [7], but more recently, experiments are being conducted to create speech not only from written text but also from articulatory movement, or even from brain signals. Procedures in which the input for speech synthesis is some representation of human articulation and the output is audible speech are called Articulatory-to-Acoustic Mapping (AAM) [7]. On the contrary, Acoustic-to-Articulation Inversion (AAI) is aimed at obtaining articulatory movements from speech. Brain-Computer Interfaces (BCI) are devices that process brain activity and send signals to external software, allowing a user to control devices with their thoughts.

In his own individual research and in the research groups he led, Tamás Gábor Csapó focused on speech synthesis. He was particularly concerned with the question of how to develop a useful and easily accessible application for the population with speech impediments, either by utilizing articulatory recordings (UTI), and later through the detection of brain signals (EEG). It is not possible to provide a full presentation of Csapó's application-oriented research here; we briefly summarize some investigations and initiatives that we consider most significant.

5.1 Articulatory-to-Acoustic Mapping in Silent Speech Interface

In the so-called Silent Speech Interface (SSI) systems, the silent movement of the articulatory organs must be recorded using some suitable technique (this can be UTI, EMA, or other suitable methodology), and then the machine system synthesizes speech from this recording. The significance of such a technology lies in the fact that the user of the device does not have to be

able to produce sound. The most frequently mentioned example of users is people with speech impediments whose articulation is intact but who are unable to produce voice for some reason (e.g., due to total or partial removal of the larynx). Communication in noisy environments is also usually referred to, as well as message transmission in situations where the speaker must be silent but cannot communicate in writing (besides some everyday situations, military or secret service actions can also be mentioned here).

In the first Hungarian-language ultrasound-based SSI experiment [11], approximately 200 synchronized speech and UTI recordings from one speaker were utilized. From the acoustic signal, fundamental frequency and spectral parameters were extracted. Then, DNN-based machine learning was applied, where the input was the ultrasound and the output was the speech spectral parameters. Using the f_0 parameters from the original speech and the spectral parameters estimated by machine learning, the authors synthesized sentences in which words, and sometimes even entire sentences, became intelligible.

The acoustic structure produced from ultrasound recordings is naturally suitable primarily for estimating spectral characteristics, as the tongue is one of the main contributors in supraglottal articulation. However, according to previous experiments, estimating the fundamental frequency is also possible with the help of images recorded by ultrasound. This is based on the fact that voicing is related to articulatory characteristics. In their experiments, Grósz and colleagues [26] trained two DNNs in parallel, one for voicing (voiced vs. voiceless), and one for estimating the specific fundamental frequency values and spectral parameters. They obtained a correlation coefficient of 0.74 between the estimated and original fundamental frequency curve. In the subjective listening tests, the participants rated the sentences synthesized with the original and the predicted f_0 curve as nearly equally natural. Further experiments concerning f_0 were also conducted, for example, using Wave-Glow, which is one of the most recently developed neural vocoders. Its advantage is that it is relatively simple but fast. Based on the perceptual tests, in the case of three out of four speakers, natural-sounding synthesized speech was successfully produced using Wave-Glow in the articulatory-to-acoustic conversion based on ultrasound recordings [15].

In one study concerning SSI, researchers investigated how variability between different recordings (session dependency) affects the success of articulatory-to-acoustic conversion in the ultrasound-based speech synthesis [21]. Based on previous research results, the authors expected that using recordings where the ultrasound probe was removed and then re-placed under the speaker's chin would significantly degrade the accuracy of the pre-trained system. They found that the synthesized speech was unintelligible when the system was trained on one recording and evaluated on another (without adaptation). Even with the same recording, there were large differences in the performance of the DNN models, in a session-dependent manner. Finally, the DNN model trained on the specific recording performed better with DNN adaptation, and furthermore, DNN adaptation also led to faster operation.

5.2 Acoustic-to-articulation inversion

Developments have also begun in the opposite direction, i.e., in the field of acoustic-to-articulation inversion. Such systems can be utilized, for example, in computer-assisted language learning, where the pronunciation of individual sounds can be practiced based on the visualization of the ultrasound recording. In this research, Tamás Gábor Csapó and his colleagues also applied DNN-based machine learning: this time, the spectral parameters of speech formed the input, and the output was the generated ultrasound video (e.g., [38]).

5.3 Brain-computer interface

The significance of Tamás Gábor Csapó's initiative aimed at the domestic development of communication brain-computer interface (BCI) is also beyond dispute. BCI applications can enable

natural or near-natural quality communication for individuals who are unable to produce audible speech due to physical or neurological damage. Measured neural activity serves as the basis for real-time speech synthesis. Csapó and his colleagues attempted to synthesize speech from brain activity recorded using electroencephalography (EEG) (e.g., [1]); and also to produce ultrasound images from EEG signals [8]. In 2023, Tamás Gábor Csapó organized a workshop on this topic (<https://smartlab.tmit.bme.hu/moonshine-2023/>) to foster international connections for the project that was already under development at the time.

6 Ongoing projects and future plans

In his most recent project, Csapó proposed a Brain-Computer-Interface (BCI)—a multimodal approach that integrates the analysis of cognitive processes in the brain, articulatory kinematics, and the resulting speech signal (2022–2026: *Analysis of articulation and brain signals for speech-based brain-computer interfaces*; NKFIH FK-22 research project; PI: Tamás Gábor Csapó, Dávid Sztahó, Budapest University of Technology and Economics). The research group conducts multimodal analysis and synthesis experiments using EEG, UTI, and speech data. One for the key objectives is to predict ultrasound tongue images from the EEG signals using a fully connected deep neural network.

Building on the foundations laid through previous collaborative work with Tamás Gábor Csapó, in 2024, a new research group was established for the *Experimental Analyses of Phonological Contrasts in Typical and Atypical Speakers* (MTA–HUN-REN NYTK Lendület „Momentum” Neurophonetics Research Group; PI: Tekla Etelka Grácz, HUN-REN Hungarian Research Centre for Linguistics). This project investigates the acoustic and articulatory realisation of specific phonological contrasts—both segmental and prosodic—in neurotypical and neuroatypical speaker groups (incl. people from three speaker groups with neuroatypical backgrounds of basically different conditions and speech features: Multiple Sclerosis, Parkinson’s Disease and Down Syndrome). UTI is planned to be used in the analysis of (i) vowel contrasts (examining whether tongue position varies as a function of vowel height in the neuroatypical vs. control groups); (ii) voicing contrast of obstruents (assessing whether articulatory manoeuvres differ between neuroatypical and neurotypical speakers); and (iii) place of articulation in consonants (studying whether and how speech difficulties in neuroatypical groups lead to a reduction in place-of-articulation distinctions). Through these studies, UTI is expected to contribute to a deeper understanding of cognitive and articulatory mechanisms underlying neuroatypical speech.

7 Acknowledgments

The preparation of the paper was supported by the TKA-DAAD Scholarship with number 1773751 (A. D., K. J.), the EKÖP-24 grant (K. J.), and the Bolyai János Research Scholarship of the Hungarian Academy of Sciences (A. D.).

8 References

- [1] ARTHUR, F. V., and T. G. CSAPÓ: *Deep learning alapú agyi jel feldolgozás és beszédszintézis előkészítő munkálatai*, In: G. BEREND, G. GOSZTOLYA, and V. VINCZE (eds.): *XVIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2022)*, pp. 185–198. Szegedi Tudományegyetem Informatikai Intézet, Szeged, 2022.
- [2] Articulate Instruments Ltd.: *Ultrasound Stabilisation Headset Users Manual: Revision 1.4*. Articulate Instruments Ltd. Edinburgh, 2008.
- [3] Articulate Instruments Ltd.: *Articulate Assistant Advanced User Guide: Version 2.14*. Articulate Instruments Ltd., Edinburgh, 2012.
- [4] BEŇUŠ, S., and A. I. GAFOS: *Articulatory characteristics of Hungarian ‘transparent’ vowels*, *Journal of Phonetics*, 35, pp. 271–300, 2007. <https://doi.org/10.1016/j.wocn.2006.11.002>

- [5] BOLLA, K.: *A magyar hosszú mássalhangzók képzése (Kinoröntgenografikus vizsgálat számítógéppel)*, Magyar Fonetikai Füzetek, 7, pp. 7–55, 1981.
- [6] BOLLA, K.: *A magyar magánhangzók és rövid mássalhangzók képzési sajátosságainak dinamikus kinoröntgenográfiai elemzése*, Magyar Fonetikai Füzetek, 8, pp. 5–62, 1981.
- [7] CSAPÓ, T. G.: *A nyelvmozgás ultrahangos vizsgálata és az automatikus elemzés alkalmazási lehetőségei a beszédtechnológiában*. In: A. DEME & Á. KUNA (eds.): *Tanulmányok a nyelvészet alkalmazásainak területéről*, pp. 197–219, ELTE Eötvös Kiadó, Budapest, 2022.
- [8] CSAPÓ, T. G., F. V. ARTHUR, P. NAGY, and Á. BONCZ: *Towards Ultrasound Tongue Image prediction from EEG during speech production*, In: *Proceedings of the 24th International Speech Communication Association, INTERSPEECH 2023*, pp. 1164–1168. International Speech Communication Association (ISCA), Dublin, 2023. <https://doi.org/10.21437/Interspeech.2023-40>
- [9] CSAPÓ, T. G., and D. CSOPOR: *Ultrahangos nyelvkontúrkövetés automatikusan: a mély neuronhálókra alapuló AutoTrace eljárás vizsgálata*, *Beszédkutató* 2015, pp. 177–187, 2015.
- [10] CSAPÓ, T. G., A. DEME, T. E. GRÁCZI, A. MARKÓ, and G. VARJASI: *Szinkronizált beszéd- és nyelvultrahang-felvételek a SonoSpeech rendszerrel*. In: V. VINCZE (ed.): *XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2017)*, pp. 339–346, Szegedi Tudományegyetem Informatikai Intézet, Szeged, 2017. <http://rgai.inf.u-szeged.hu/project/mszny2017/files/kotet.pdf>
- [11] CSAPÓ, T. G., T. GRÓSZ, T., L. TÓTH, and A. MARKÓ, A.: *Beszéd-szintézis ultrahangos artikulációs felvételekből mély neuronhálók segítségével*. In: V. VINCZE (ed.): *XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2017)* (pp. 181–192). Szegedi Tudományegyetem Informatikai Intézet, Szeged, 2017. <https://rgai.inf.u-szeged.hu/sites/rgai.sed.hu/files/kotet.pdf>
- [12] CSAPÓ, T. G., and S. M. LULICH: *Error Analysis of Extracted Tongue Contours from 2D Ultrasound Images*. In: *Proceedings of Interspeech 2015*, pp. 2157–2161. ISCA, Dresden, 2015. <https://doi.org/10.21437/Interspeech.2015-486>
- [13] CSAPÓ, T. G., and K. XU: *Quantification of Transducer Misalignment in Ultrasound Tongue Imaging*. *Proc. Interspeech 2020*, pp. 3735–3739, Interspeech, Sanghaj, 2020. <https://doi.org/10.21437/Interspeech.2020-1672>
- [14] CSAPÓ, T. G., K. XU, A. DEME, T. E. GRÁCZI, and A. MARKÓ: *Transducer Misalignment in Ultrasound Tongue Imaging*. In: M. TIEDE, D. H. WHALEN, and V. GRACCO (eds.): *Proceedings of the 12th International Seminar on Speech Production*, pp. 166–169, Haskins Press, New Haven (CT), 2021.
- [15] CSAPÓ, T. G., Cs. ZAINKÓ, L. TÓTH, G. GOSZTOLYA, and A. MARKÓ: *Ultrasound-Based Articulatory-to-Acoustic Mapping with WaveGlow Speech Synthesis*. In: H. MENG, B. XU, and T. ZHENG (eds.): *Interspeech 2020*, pp. 2727–2731, International Speech Communication Association (ISCA), Sanghaj, 2020. <https://doi.org/10.21437/Interspeech.2020-1031>
- [16] DEME, A., R. GREISBACH, A. MARKÓ, M. MEIER, M. BARTÓK, J. JANKOVICS, and Zs. WEIDL: *Tongue and jaw movements in high-pitched soprano singing: A case study*, *Beszédkutató* 2016, pp. 121–138, 2016.
- [17] DEME, A., R. GREISBACH, M. MEIER, M. BARTÓK, J. JANKOVICS, Zs. WEIDL, and A. MARKÓ: *Tongue and jaw articulation of soprano singers at high pitch in Hungarian and German*. In: J. DANG, and A. LI (eds.): *11th International Seminar on Speech Production (ISSP 2017) Book of Abstracts*, pp. 110–113, ISSP, Tianjin, 2017.
- [18] EPSTEIN, M. A., and M. STONE: *The tongue stops here: Ultrasound imaging of the palate (L)*, *The Journal of the Acoustical Society of America*, 118(4), pp. 2128–2131, 2005.
- [19] FASEL, I., and J. BERRY: *Deep Belief Networks for Real-Time Extraction of tongue contours from ultrasound during speech*. In: *Proceedings of ICPR*, pp. 1493–1496, ICPR, Istanbul, 2010. <https://doi.org/10.1109/ICPR.2010.369>

- [20] GÓSY, M.: On the history of palatography in Hungarian phonetics. *Journal of the International Phonetic Association*, 53(3), pp. 682–693, 2023.
- [21] GOSZTOLYA, G., T. GRÓSZ, L. TÓTH, A. MARKÓ, and T. G. CSAPÓ: *Applying DNN adaptation to reduce the session dependency of ultrasound tongue imaging-based silent speech Interfaces*, *Acta Polytechnica Hungarica*, 17(7), pp. 109–124, 2020. <https://doi.org/10.12700/APH.17.7.2020.7.6>
- [22] GRÁCZI, T. E., T. G. CSAPÓ, M. BARTÓK, A. DEME, and A. MARKÓ: *The realization of voicing opposition in alveolar fricatives in Hungarian: Preliminary study on articulation and acoustics*, *Beszédtudomány / Speech Science 1*, pp. 22–56, 2020.
- [23] GRÁCZI, T. E., T. G. CSAPÓ, M. BARTÓK, A. DEME, and A. MARKÓ: *Articulatory and acoustic differentiation of /s/ and /ʃ/ in children's speech: longitudinal case studies*. In: J. BÓNA (ed.): *(Dis)fluencies in children's speech*. Akadémiai Kiadó, Budapest, 2021. https://mersz.hu/dokumentum/m873dfics_35/ <https://doi.org/10.1556/9789634547099.4>
- [24] GRÁCZI, T. E., T. G. CSAPÓ, A. DEME, K. JUHÁSZ, and A. MARKÓ: *A réshangok zöngésségével összefüggő nyelvpozíciós jellemzők a megelőző magánhangzóban*, *Nyelvtudományi Közlemények*, 116, pp. 155–190, 2020.
- [25] GRÁCZI, T. E., T. G. CSAPÓ, A. DEME, K. JUHÁSZ, and A. MARKÓ: *Tongue root position in VC sequences with regard to the phonetic realization of obstruent voicing: A preliminary study on Hungarian*. In: M. TIEDE, D. H. WHALEN, and V. GRACCO (eds.): *Proceedings of the 12th International Seminar on Speech Production*, pp. 198–201, Haskins Press, New Haven (CT), 2021.
- [26] GRÓSZ, T., L. TÓTH, G. GOSZTOLYA, T. G. CSAPÓ, and A. MARKÓ: *Kísérletek az alapfrekvencia becslésére mély neuronhálós, ultrahang-alapú némabeszéd-interfészekben*. In: V. VINCZE (ed.): *XIV. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2018)*, pp. 196–2005, Szegedi Tudományegyetem Informatikai Intézet, Szeged, 2018.
- [27] KELSEY, C. A., F. D. MINIFIE, and T. J. HIXON: *Applications of ultrasound in speech research*, *Journal of Speech and Hearing Research*, 12, 564–575, 1969. <https://doi.org/10.1044/jshr.1203.564>
- [28] KOCHETOV, A.: *Research methods in articulatory phonetics I: Introduction and studying oral gestures*, *Language and Linguistics Compass*, 14(1), pp. 1–29, 2020. <https://doi.org/10.1111/lnc3.12368>
- [29] LI, M., Ch. KAMBHAMETTU, and M. STONE: *Automatic contour tracking in ultrasound images*, *Clinical Linguistics and Phonetics*, 19(6–7), pp. 545–554, 2005. <https://doi.org/10.1080/02699200500113616>
- [30] LOTZ, J.: *Egy magyar röntgen-hangosfilm és néhány fonológiai kérdés*, *Magyar Nyelv*, 62, pp. 257–266, 1966.
- [31] MÁDY, K.: *Magyar magánhangzók vizsgálata elektromágneses artikulográffal normál és gyors beszédben*, *Beszédkutatás 2008*, pp. 52–66, 2008.
- [32] MARKÓ, A., M. BARTÓK, T. G. CSAPÓ, T. E. GRÁCZI, and A. DEME: *Az /i:/ artikulációs és akusztikai sajátosságai harmonikusan és antiharmonikusan toldalékolódó tövekben*, *Nyelvtudományi Közlemények*, 115, pp. 233–254, 2019.
- [33] MARKÓ, A., T. E. GRÁCZI, K. JUHÁSZ, and A. DEME: *A brief historical survey of articulatory studies on the Hungarian language*. In: J. BÓNA (ed.): *Proc. Sixth International Workshop on the History of Speech Communication Research (HSCR 2024)*, pp. 106–115, 2024.
- [34] PALO, P., S. M. LULICH, and D. AALTO: *Interaction of phonetic materials and simulated probe position on the mean square error metric of tongue ultrasound probe alignment – a methodological case study*, *Beszédtudomány – Speech Science 2025(1)*, pp. 74–89, 2025. <https://doi.org/10.15775/Besztud.2025.1.74-89>
- [35] PERCIVAL, M., T. G. CSAPÓ, M. BARTÓK, A. DEME, T. E. GRÁCZI, and A. MARKÓ: *Tongue root and voicing in Hungarian singleton and geminate obstruents*. In: *12th International*

- Seminar on Speech Production*, 2020. https://issp2020.yale.edu/S08/percival_08_15_156_abstract.pdf
- [36] PERCIVAL, M., T. G. CSAPÓ, A. MARKÓ, A. DEME, T. E. GRÁCZI, and M. BARTÓK: *Gemination as fortition? Articulatory data from Hungarian*. Presentation at LabPhon17 virtual conference. Vancouver, July 6-8, 2020, 2020. https://labphon.org/sites/default/files/previous_conferences/LP17/abstracts/LabPhon_17_paper_308.pdf
- [37] PERCIVAL, M., T. G. CSAPÓ, M. BARTÓK, A. DEME, T. E. GRÁCZI, and A. MARKÓ: *Ultrasound imaging of Hungarian geminates*. Presentation at UltraFest IX. Ultrasound Imaging for Speech and Language. Virtual Conference. October 21-24, 2020. 2020. https://ultrafest2020.indiana.edu/abstracts/UltraFest_IX_Percival_Csapo_Bartok_Deme_Graczi_Marko_Hungarian.pdf.
- [38] PORRAS, D., A. SEPULVEDA-SEPULVEDA, and T. G. CSAPÓ: *DNN-Based Acoustic-to-Articulatory Inversion Using Ultrasound Tongue Imaging*. In: *Proceedings of the International Joint Conference on Neural Networks*, pp. 1–8, IEEE, New Jersey, 2019. <https://doi.org/10.1109/IJCNN.2019.8851769>
- [39] RECASENS, D., and C. RODRÍGUEZ: *A study on coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound*, *Journal of Phonetics*, 59, pp. 58–75, 2016. <https://doi.org/10.1016/j.wocn.2016.09.002>
- [40] SCOBIE, J. M., E. LAWSON, S. COWEN, J. CLELAND, and A. A. WRENCH: *A common coordinate system for mid-sagittal articulatory measurement*, *Working Paper WP-20. June 2011*. Queen Margaret University, Edinburgh, 2011. <https://core.ac.uk/reader/161924590>
- [41] STONE, M.: *A guide to analysing tongue motion from ultrasound images*, *Clinical Linguistics & Phonetics*, 19(6–7), pp. 455–501, 2005. <https://doi.org/10.1080/02699200500113558>.
- [42] STONE, M.: *Laboratory Techniques for Investigating Speech Articulation*. In: W. J. HARDCASTLE, J. LAVER, and F. GIBBON (eds.): *The Handbook of Phonetic Sciences*, pp. 9–38, Willey-Blackwell, Oxford, 2010. <https://doi.org/10.1002/9781444317251.ch1>
- [43] SZENDE, T.: *A magyar hangrendszer néhány összefüggése röntgenográfiai vizsgálatok tükrében*, *Magyar Nyelv*, 70, pp. 68–77, 1974.
- [44] TANG, L., T. BRESSMANN, and G. HAMARNEH: *Tongue contour tracking in dynamic ultrasound via higher-order MRFs and efficient fusion moves*, *Medical Image Analysis*, 16(8), pp. 1503–1520, 2012. <https://doi.org/10.1016/j.media.2012.07.001>
- [45] TRENCSENYI, R.: *A nyelvkontúrkövető algoritmusok és a gépi tanulás összekapcsolhatóságának vizsgálata*. In: G. BEREND, G. GOSZTOLYA, and V. VINCZE (eds.): *XVI. Magyar Számítógépes Nyelvészeti Konferencia: MSZNY 2020. Szeged, 2020. január 23–24.*, pp. 233–244, Szegedi Tudományegyetem Informatikai Intézet, Szeged, 2020.
- [46] TRENCSENYI, R.: *MRI- és UH-felvételek geometriai elemzése a beszédszintézisben*, *Acta Medicinæ et Sociologica*, 11(31), pp. 55–65, 2020.
- [47] TRENCSENYI, R., and L. CZAP: *Possible Methods for Combining Tongue Contours of Dynamic MRI and Ultrasound Records*, *Acta Polytechnica Hungarica*, 18(4), pp. 143–160, 2021.
- [48] WHALEN, D. H., J. KANG, R. IWASAKI, G. SHEJAAYA, B. KIM, K. D. ROON, and M. K. TIEDE: *Accuracy assessments of hand and automatic measurements of ultrasound images of the tongue*. In: S. CALHOUN, P. ESCUDERO, M. TABAIN, and P. WARREN (eds.): *Proceedings of the 19th International Congress of Phonetic Sciences*, pp. 542–546, Australasian Speech Science and Technology Association Inc., Canberra, 2019.
- [49] ZHARKOVA, N., F. E. GIBBON, and W. J. HARDCASTLE: *Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation*, *Clinical Linguistics and Phonetics*, 29, pp. 249–265, 2015. <https://doi.org/10.3109/02699206.2015.1007528>
- [50] ZHARKOVA, N., and N. HEWLETT: *Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /a/*, *Journal of Phonetics*, 37, pp. 248–256, 2009. <https://doi.org/10.1016/j.wocn.2008.10.005>

CONTEXTUALISING THE WAX CYLINDER RECORDINGS OF THE INSTITUTION NATIONALE DES SOURDS-MUETS DE PARIS

Sarah Fuchs¹, Apolline Gouzi²

¹ Royal College of Music, London, United Kingdom,

² University of Cambridge, United Kingdom

sarah.fuchs@rcm.ac.uk

Abstract: Among the many thousands of analogue sound recordings preserved at the Bibliothèque Nationale de France are 228 unpublished wax cylinder recordings that once belonged to the Institution Nationale des Sourds-Muets de Paris (now known as the Institut National de Jeunes Sourds de Paris). Recently digitised and descriptively catalogued by the BNF, these recordings attest to the extraordinary array of activities that took place in the Institution’s Laboratoire de la Parole over the first few decades of the twentieth century on the initiative of its first director, Hector Marichelle (1862–1929). Based on the co-authors’ ongoing research related to these recordings, this article takes an initial step towards contextualising the collection as a whole. It describes the current state of the collection and its cataloguing system, and lays out what is known about when, how, and why the recordings were initially made. Taking selected cylinders as case-studies, it discusses some of the practical and methodological challenges the co-authors have encountered in the course of their research, in particular as they have attempted to create text-based transcriptions of the recordings. This article thus aims to contribute to the intertwined histories of experimental phonetics and deaf pedagogy in early twentieth-century Europe, as well as to larger interdisciplinary conversations about how sensitive sonic heritage has become the stuff of scholarly study.

1 Introduction

The Bibliothèque Nationale de France’s Département Son, Vidéo, Multimédia currently holds in its collection 228 wax cylinder recordings that once belonged to the Institution Nationale des Sourds-Muets de Paris [1]. Recently digitised and descriptively catalogued by the BNF, the recordings in this collection attest to the extraordinary array of activities that took place in the Institution’s Laboratoire de la Parole over the first few decades of the twentieth century on the initiative of its first director, Hector Marichelle (1862–1929). The sheer variety of the voices captured on these cylinders is one of the most striking things about the collection: recordings of the Institution’s deaf students jostle noisily with those made by scientists, statesmen, and singers. Very little has been written about the circumstances under which the cylinders were recorded, still less those under which they became a part of the Bibliothèque Nationale de France’s collection. Indeed, much of what we have learned about the collection has come from our work with the cylinders themselves, hence we begin with our observations about the current state of the collection and its cataloguing system.

2 Current state of the collection and cataloguing system

The collection would seem have been absorbed by one of the Département Son, Vidéo, Multimédia’s institutional predecessors at some point over the mid-twentieth century. Tracing this institutional history—from the Archives de la Parole (founded in 1911) to the Musée de la Parole et du Geste (1927–28) to the Phonothèque Nationale (1938) to the Département de l’Audiovisuel (1994) to today’s Département Son Vidéo, Multimédia—is beyond the scope of the present study, all the more so because any such history would need to take account of how

these institutions intersected with each other (and with other institutions, such as the Institut de Phonétique, with which both the Musée de la Parole et du Geste and the Phonothèque Nationale shared space at 19 rue des Bernardins [2]). Even if we were able to pin down which one of these institutions acquired the collection, many questions would remain unanswered.

The collection consists of 228 wax cylinders, many of which are preserved in their original cylindrical cardboard boxes. Most of the boxes and/or box lids have handwriting on them, sometimes written in more than one hand and/or with more than one kind of writing instrument. For example, in the centre of the box lid for the cylinder now catalogued as CYL-572 is a description of the recording session: ‘J[eudi?] 18 d[écembre?] [19]24 / El[ève?] de M. Janin’ (‘Thursday 18 December 1924 / Student of Mr. Janin’). Written above it (in a different hand and with a different writing instrument) is the number ‘191’, which appears to be a remnant of a now defunct numbering system. It is impossible to establish exactly when either of these inscriptions was added to the box lid, but it is plausible that the first dates from around the same time as the recording session and the second from some later stage.

Curled up inside about fifty of the boxes are additional handwritten notes that explain the contents of the cylinders in more detail. Inside the box for CYL-572, for example, is a yellowed piece of graph paper on which is written the full date of recording (‘18 December 1924’), the three diction exercises around which the recording session revolved (‘a, pa, and i’), and the names of the people involved in the recording (Institution Nationale des Sourds-Muets Professor Janin and five of his deaf and hard-of-hearing students). Two of the boxes contain spools of paper on which are reproduced the fine wavy lines etched into the wax cylinders throughout the recording session—in other words, graphic analyses of speech sounds. It is worth pointing out in passing that some of these notes would seem to have become separated from their cylinders, and thus are occasionally found in unexpected boxes.

Today, the entire collection—cylinders, boxes and box lids, notes, and graphic analyses—resides in a storeroom in the basement of François-Mitterrand site of the Bibliothèque Nationale de France, where the cylinders are shelved according to their *cotes* (catalogue numbers). The catalogue numbers appear to have been assigned to the cylinders more or less arbitrarily, as the cataloguer came to them. The numbering system is mostly consecutive, with the majority of the cylinders assigned catalogue numbers between CYL-455 and CYL-669, which suggests that the cataloguer encountered the cylinders as an almost complete collection from the start. However, a few cylinders would seem to have been integrated into this collection at a slightly later stage, or so we have assumed on the basis of several noticeable gaps in the numbering system (for example, the next catalogue number after CYL-669 is CYL-693).

Unlike its catalogue numbering system, the BNF’s descriptive catalogue imposes a very particular kind of order on the collection. Each of the cylinders has been classified according to one of four categories: ‘essais techniques’ (‘technical tests’), ‘exercices de diction’ (‘diction exercises’), ‘enregistrements de paroles’ (‘spoken-word recordings’), and ‘enregistrements musicaux’ (‘musical recordings’). Spoken-word recordings and musical recordings are further classified according to one of several sub-categories (mostly orientated around genre, though there are also a few recordings featuring speakers of foreign languages, including French dialects). Within each category and sub-category, the cylinders are listed in ascending order according to their catalogue numbers, but there is no way to view a complete list of the cylinders outside these categories.

Taken together, the four main categories classify the cylinders according to their apparent function (i.e. testing the laboratory’s equipment and documenting diction exercises) or the primary mode of expression featured in the recording (i.e. speaking or singing). Although some of the recordings fall fairly neatly into one or another of these categories, others are far less straightforward. The descriptive catalogue provides no information about how the categorisa-

tion system was devised, nor how the cylinders were classified. It does, however, furnish detailed descriptions of what the cataloguer heard in each of the recordings, as well as partial transcriptions of the handwriting on the boxes, box lids, and notes inside the boxes.

The descriptive catalogue has its home in the BNF's online database 'Archives et Manuscrits', which fulfils a different purpose from its Catalogue Général (and with which it is not consistently cross-referenced). The digital transfers of the cylinders are destined for the BNF's digital library Gallica, where one imagines they may be organised differently, but in the meantime one must identify a cylinder in the descriptive catalogue, then make direct contact with the staff of the Département Son, Vidéo, Multimédia to arrange a listening session. The descriptive catalogue's organisation system thus shapes one's understanding of the cylinders before they have even been heard.

3 Transcribing the cylinders

Listening carefully to the cylinders has played an important part of the present project, one strand of which has involved creating text-based transcriptions of the entire collection. One of the first decisions we made was to listen to the cylinders in ascending order of catalogue number rather than by category, in an effort to come to our own understanding of the collection's logic (or lack thereof). Almost immediately, it became clear that there was considerable overlap between the categories, both in terms of the cylinders' apparent functions and the modes of expression captured on them. Singing, for example, can be heard across several different categories, including 'essais techniques' ('technical tests') and 'exercices de diction' ('diction exercises') as well as 'enregistrements musicaux' ('musical recordings'), where it would seem to have been understood as fulfilling different functions.

On one level, singing was simply one of several ways of testing that the equipment was working properly. So it is that one can hear a technician singing fragments from same popular song (the 'Polka des Englishs' [*sic*], popularised in the early twentieth century by Charlus and other *café-concert* singers) in several different cylinders, including not only some of those classified as 'essais techniques' (for example, CYL-609 and CYL-632) but also those classified among the 'exercices de diction' (such as CYL-606, which might just easily have been classified among the 'enregistrements des paroles' because it also features a recitation of Jean de La Fontaine's fable 'Le corbeau et le renard') and the 'enregistrements musicaux' (CYL-617, which also contains a performance of Alfred de Vigny's 'Le cor').

Singing was also a way of teaching and testing diction. In many different cylinders—including not only those made by Institution Nationale des Sourds-Muets professors and pupils but also by professional singers—one hears pitch-pipes used to establish the pitch at which diction exercises are to be executed. In the aforementioned CYL-572, for example, the students (who are described in an accompanying note as having differing degrees of hearing loss) attempt to match the pitch given by the pitch-pipe (see Table 1). A pitch-pipe similarly sounds at the outset of CYL-498, after which an opera singer identified as Ketty Lapeyrette (1884–1960) intones the vowel-sounds 'A - O - É - E - OU - I' in tune with the pitch-pipe, then runs through some scales, speaks the words 'Bonheur et gloire aux défenseurs de Verdun' ('Good luck and glory to the defenders of Verdun'), and finally sings an excerpt from 'Mon cœur s'ouvre à ta voix' from Saint-Saëns's 1877 *grand opéra Samson et Dalila*, which remained in the French operatic repertoire well into the twentieth century (see Table 2).

Across the collection, the line separating diction exercises and singing exercises is sometimes very blurry indeed, a reminder of the widely shared understanding of the French language as inherently musical [3, 4] and how much proper pronunciation mattered to the French school of singing [5]. In some instances, we have been unable to establish exactly what function singing fulfilled in the context of the recording session, even among cylinders classified as 'enregistrements musicaux'. CYL-498 (discussed above) is one such cylinder. It would seem to have

Table 1. Transcription of CYL-572

<p>Diapason [Élève 1] : A - A [Élève 2] : A [Élève 3] : A [Élève 4] : A [Élève 5] : A</p> <p>Diapason x2 [Élève 1] : A [Élève 2, crié] : PA [Élève 3, tenu] : PA [Élève 4] : PA [Élève 5] : PA</p> <p>Diapason x2 [Élève 1] : I [Élève 2] : I [Élève 3] : I [Élève 4] : PI [Élève 5] : I Diapason</p>	<p>Pitch-pipe [Student 1] : A – A [Student 2] : A [Student 3] : A [Student 4] : A [Student 5] : A</p> <p>Pitch-pipe x2 [Student 1] : A [Student 2, shouted] : PA [Student 3, sustained] : PA [Student 4] : PA [Student 5] : PA</p> <p>Pitch-pipe x2 [Student 1] : I [Student 2] : I [Student 3] : I [Student 4] : PI [Student 5] : I Pitch-pipe</p>
---	---

Table 2. Transcription of CYL-498

<p>Diapason</p> <p>[Mlle Lapeyrette, vocalisé] : A - O - É - E - OU - I [sur le ton du diapason] A (gamme ascendante et descendante sur Do M) I (gamme ascendante et descendante sur Do M) U (gamme ascendante et descendante sur Do M) A (gamme ascendante et descendante sur Do M)</p> <p>[Mlle Lapeyrette, parlé] : Bonheur et gloire aux défenseurs de Verdun !</p> <p>[Mlle Lapeyrette, chanté] : Mon cœur s'ouvre à ta voix Comme s'ouvrent les fleurs Aux baisers de l'aurore ! Mais toi, mon bien-aimé Pour mieux sécher mes pleurs Que ta voix parle encore ! Dis moi qu'à Dalila Tu reviens pour jamais. Redis à ma tendresse Les serments d'autrefois !</p>	<p>Pitch-pipe</p> <p>[Miss Lapeyrette, vocalised]: A - O - É - E - OU - I [on the pitch of the pitch-pipe] A [ascending and descending scale on C Major] I [ascending and descending scale on C Major] U [ascending and descending scale on C Major] A [ascending and descending scale on C Major]</p> <p>[Miss Lapeyrette, spoken]: Bonheur et gloire aux défenseurs de Verdun!</p> <p>[Miss Lapeyrette, sung]: Mon cœur s'ouvre à ta voix Comme s'ouvrent les fleurs Aux baisers de l'aurore! Mais toi, mon bien-aimé Pour mieux sécher mes pleurs Que ta voix parle encore! Dis moi qu'à Dalila Tu reviens pour jamais. Redis à ma tendresse Les serments d'autrefois!</p>
--	---

been classified among the ‘enregistrements musicaux’ because the cataloguer heard Lapeyrette’s performance of ‘Mon cœur s’ouvre à ta voix’ as the main event of the recording session. But what if it was not intended to be a performance? It is possible that Lapeyrette’s pitched intonation of the vowel-sounds ‘A-O-É-E-OU-I’—and/or, indeed, her singing of ‘Mon cœur s’ouvre à ta voix’—was some kind of test, whether of her own diction or of the equipment. It is possible, in other words, that this cylinder belongs not in the category ‘enregistrements musicaux’ but instead in ‘exercices de diction’ or ‘essais techniques’. The same may be true of many other cylinders in the collection, though further research is required before we can say this with any certainty.

As the examples cited above illustrate, the cylinders in this collection capture an array of activities, often more than one in the same recording session. Throughout our transcription process, we have aimed to transcribe not just what one might think of as the ‘content’ of each cylinder, but also the different modes of expression with which that content is communicated, leaving aside for the moment the larger issue of how the cylinders should be categorised (if at all). In practical terms, this has involved identifying the voice we hear at any given moment—by name whenever possible, though in most instances we have had to settle for something far more generic—then describing what that voice is doing (or attempting to do): speaking, shouting, singing, vocalising, intoning, whispering, whistling, humming, clicking, laughing.

Describing voices in this way is inevitably subjective, dependent as it is on our own expectations and experiences, not to mention our own highly specific (and, as professional musicians, highly specialised) senses of hearing. We are aware that what we hear may not be what someone in another discipline—deaf education, phonetics, speech therapy—will hear, and thus we expect that our transcription system will continue to evolve as the project itself becomes more interdisciplinary in nature. We also anticipate that the archival record, which we are still actively exploring, will provide answers to at least some of our lingering questions.

4 Listening (and not listening) in the Laboratoire de la Parole

How these cylinders were meant to be listened to in their own time (let alone by researchers a hundred years later) is a question only partially answered by that archival record, or at least by the parts of it we have so far been able to consult. What we currently know about the Laboratoire de la Parole’s activities comes from several sources, including the archives of the Institut National de Jeunes Sourds de Paris [6] and those of the Institut de Phonétique, to which the Laboratoire de la Parole was annexed in the 1920s [7], as well as a handful of newspaper articles [8, 9, 10, 11]. Alongside these materials, we have also read Marichelle’s published writings, which shed light on the methods and machines used in the Laboratoire de la Parole [12, 13, 14].

Installed in the Institution Nationale des Sourds-Muets de Paris in 1912, the Laboratoire de la Parole aimed to document the Institution’s teacher-training programme, to track and observe deaf students’ linguistic progress through graphic means, and to pursue research related to experimental phonetics [15, 16]. Taken as a whole, such work was believed to have wide relevance: according to one of the Laboratoire’s early advertisements, ‘Les résultats sont applicables à l’Enseignement où à la Rééducation de la Parole; ils intéressent les méthodes de *Lecture*, de *Diction*, de *Déclamation* et de *Chant*, la correction des *Vices de Prononciation*, l’enseignement des *Langues étrangères*, l’enseignement de la *Parole aux Sourds-Muets* et aux *Aphasiques*, la culture de l’ouïe chez les sourds partiels’ (‘The results are applicable to the *Teaching or Re-education of Speech*; they are relevant to methods of *Reading, Diction, Declamation, and Singing*, to the correction of bad habits of *Pronunciation*, to the teaching of foreign *Languages*, to the teaching of *Speech to Deaf-Mutes* and people with *Aphasia*, to the culture of hearing among the partially deaf’) [17].

Throughout its early years in particular, the *Laboratoire de la Parole* attracted the attention of scientists, singers, and statesmen as well as members of the press, whose accounts of their visits made front-page news in the Parisian press. Such accounts invariably focused on the *Laboratoire*'s machines, in particular its cinematographic and phonographic devices, with which visitors would seem to have been encouraged to engage. One of these was the *Phonoscope*, which inventor Georges Demenÿ (1850–1917) had designed in the early 1890s following his and Marichelle's experimental efforts to teach deaf students to speak using a series of rapidly projected chronophotographs (it was these experiments that gave rise to the now-famous series of images spelling out 'Je vous aime', or 'I love you', shown in Figure 1 [18]. Although Demenÿ lost interest in deaf education soon after designing the *Phonoscope*, Marichelle saw the machine as one of the main methods by which the deaf could learn to communicate with the wider population, whether by lip-reading or by articulation [19, 20].

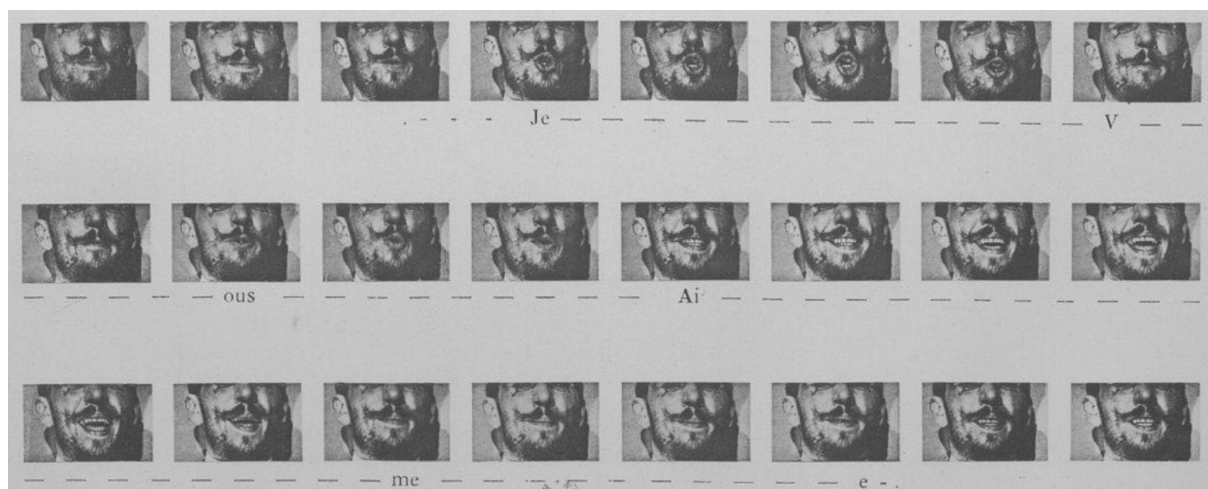


Figure 2. G. Demenÿ, *Photographie de la parole*, [p. 1], Imprimerie Lahure, Paris, Département Science et Techniques 4-TB66-80, Bibliothèque Nationale de France, Paris, gallica.bnf.fr.

Another of the *Laboratoire de la Parole*'s machines—perhaps the most important, at least for the present study—was a *Lioretgraphe Transcritteur* (so called after its inventor, Henri Lioret [1848–1938]), a specially designed system that transcribed the fine wavy lines etched into wax cylinder recordings onto long pieces of lamp-black paper, thus visualising actual vocal traces [see Figure 2].

Such visualisations facilitated the analysis of individual speech sounds, as well as timbre, intensity, and pitch quality. The latter Marichelle sometimes transferred to a musical staff, using a method he had pioneered. The poster shown below [see Figure 3] illustrates both of these visualisation methods: the top half reproduces the *Lioretgraphe*'s transcription of an excerpt from a recording made by Senator Honoré Audiffred ('Il faut encourager les travaux de Monsieur Marichelle', or 'The work of Mr. Marichelle must be encouraged'), while the bottom half shows Marichelle's rendering of the same spoken excerpt on a musical staff.

No trace survives of the visualisation Marichelle reportedly made of the voice of renowned deaf author and activist Henri Gaillard, who visited the *Laboratoire de la Parole* on 26 August 1918. In an article published a few months later in *La Gazette des Sourds-Muets*, Gaillard himself described his experience in extensive detail. He had begun by making a test recording, repeating after Marichelle the words 'toujours mieux' ('ever better'). Marichelle had then played the cylinder on the *Lioretgraphe Transcritteur*, which transferred the lines their voices had left behind in the wax to a piece of lamp-black paper. As Gaillard recalled,

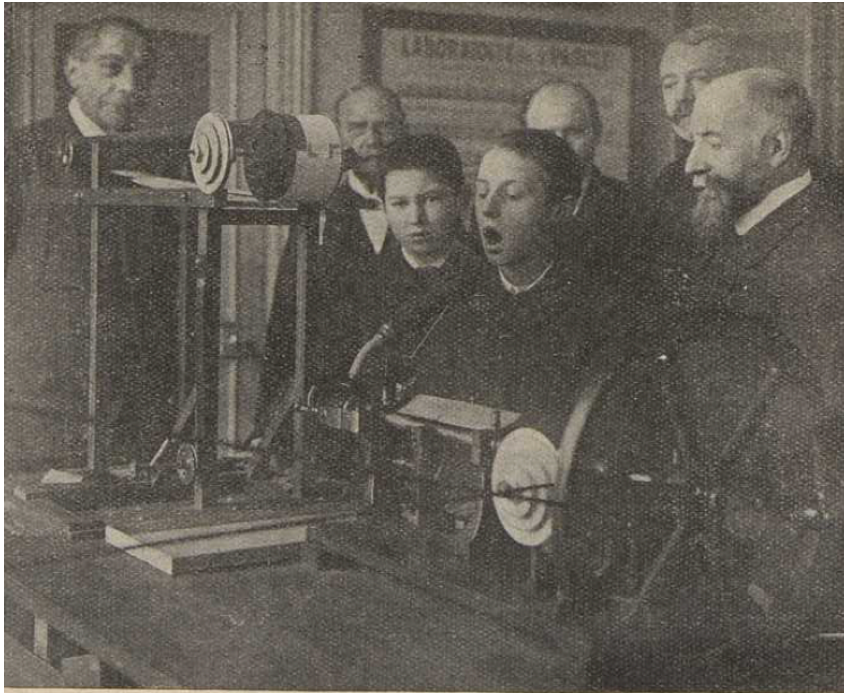


Figure 2. Photograph of Henri Lioret (left) and Hector Marichelle (right) making a recording of the voice of a deaf child, first published in *Lectures pour tous* (1918), re-published in *Le nouvel art cinématographique*, No. 2 (April 1929): [7].

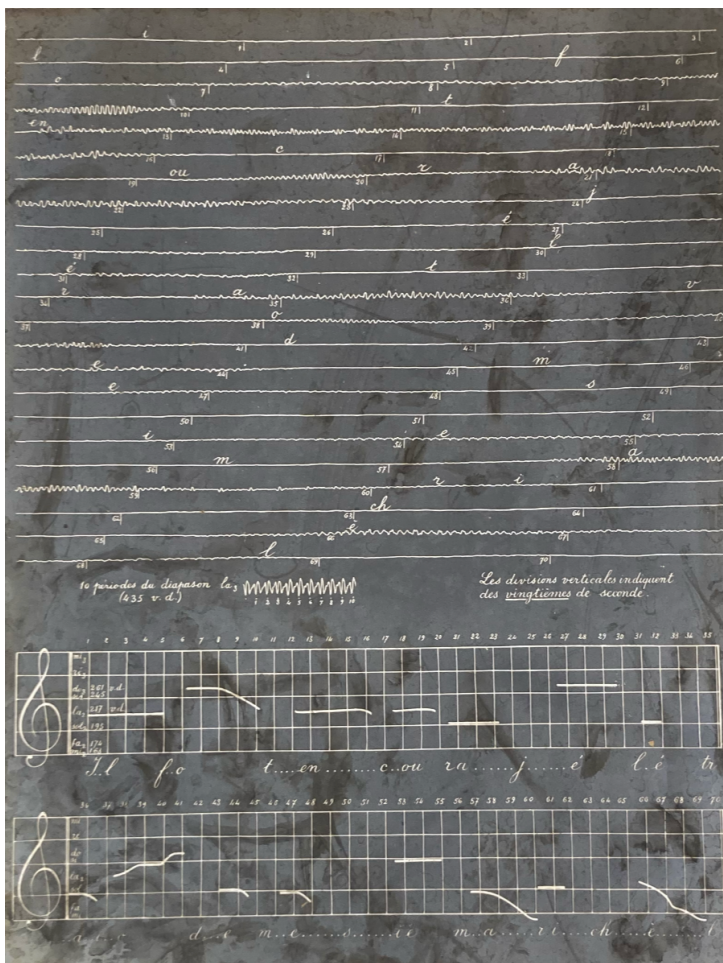


Figure 3. Etude graphique de l'intonation, affiche, pochette M-1 (œuvres diverses), Bibliothèque Patrimoniale, INJS, Paris.

‘M. Marichelle prend le rouleau où s’inscrivit ma parole et le met de côté dans un placard. Si jamais on s’avise de mettre ce cliché dans un gramophone, on pourra percevoir ma façon de prononcer. Mais M. Marichelle m’explique le graphique de ma voix. Il y a la mesure de durée [...] l’intonation [...] et la force ou intensité [...] et le timbre, c’est-à-dire la forme, le dessin de la période. Il en résulte que j’ai parcouru presque deux tons. Alors, ma voix serait donc monotone, monocorde, comme on dit’ (‘Marichelle takes the cylinder on which my speech had been inscribed and put it away in a cabinet. If anyone ever thinks to put this print in a gramophone, they will be able to hear my way of speaking. But Marichelle explains to me the graph of my voice. There is the measure of duration [...] the intonation [...] the strength or intensity [...] and the timbre, that is to say, the shape, the contour outlined. The result is that I have traversed nearly two tones. So, my voice would seem to be monotone, one-note, as they say’) [21].

As his article makes clear, Gaillard’s visit to the Laboratoire de la Parole had much more in common with that of Senator Audiffred than those of the Institution’s deaf students: he was not there to be treated, but instead to show his support of Marichelle’s work. It thus came as something of a surprise to discover a cylinder—CYL-528—matching Gaillard’s description classified among the ‘exercices de diction’ in the BNF’s descriptive finding aid, where it is in fact misidentified as an interaction between professor and pupil rather than two equals, one hearing, the other not. In identifying this error, our intention is only to illustrate just how much one’s expectations—of what a ‘deaf voice’ sounds like, and what a deaf individual could be doing in the Laboratoire de la Parole—shape what one hears, and how important it is to question such expectations when listening to sensitive sonic heritage.

5 Sensitive sonic heritage

The term ‘sensitive heritage’ is often used to describe Western collections containing non-Western artefacts, especially those acquired as a result of colonial (and other) abuses: sacred objects seized from sites of worship, for example, or the remains of enslaved people preserved for scientific purposes [22]. Over the past decade or so, certain sound recordings have begun to be described as similarly ‘sensitive’, in particular those linked to ‘the experiences of trauma, exclusion, and injustice of those whose voices were recorded, as well as the communities to which they belonged’ [23]. Many of the cylinders in this collection could be considered ‘sensitive’ in one way or another, and as such they demand special handling by their curators and their researchers.

Most sensitive of all, to be sure, are the cylinders capturing the voices of the Institution’s deaf students. In late nineteenth- and early twentieth-century France—as elsewhere in the Western world in the wake of the 1880 International Congress on Education of the Deaf—deaf children were routinely taught to communicate through speech rather than sign languages, the use of which was banned from educational programmes [24, 25]. As a formal apology issued by the 2010 International Congress on Education of the Deaf finally acknowledged, the turn towards oralism ‘led to the exclusion of Deaf citizens in educational policy and planning in most jurisdictions of the world; prevented Deaf citizens from participation in governmental planning, decision-making, and funding in areas of employment training, retraining and other aspects of career planning; hindered the abilities of Deaf citizens to succeed in various careers and [...] prevented many of them from following their own aspirations; and prevented the opportunity for many Deaf citizens to fully demonstrate their cultural and artistic contributions to the diversity of each Nation’ [26]. As evidence of the oralist method at work, above all on children who could not choose to be educated otherwise, such cylinders must be approached with extraordinary care today.

These are not the only sensitive recordings in the collection, however: others could include the cylinders recorded for the purposes of correcting provincial *patois* or curing shell-shock (including one—CYL-479—that claims to have been recorded by a ‘serg[ent] blessé’, or

‘wounded sergeant’). Taking a slightly wider view, we might consider the entire collection to be sensitive, consisting as it does of recordings that were never intended for public consumption. Examining how recordings that were implicitly private and perhaps so ephemeral—so thoroughly non-archival—that they would only ever have been listened to once and then discarded have become a part of France’s sonic heritage raises larger political questions about how digitisation and digital delivery platforms intervene in patrimonial projects across Europe. It also raises ethical questions: to what extent is it appropriate to listen to these recordings at all, let alone make them freely accessible on Gallica? Allowing these recordings to become an uncomplicated part of France’s sonic heritage, playable on command by anyone in the world, may risk betraying the communities from which they come, including those with a long and painful history of having been forced to speak only in certain ways—and, in the case of the deaf community, to speak at all.

6 Acknowledgments

Research for this paper has been supported by a Small Research Grant from the British Academy/Leverhulme Trust and a UK-France Science, Innovation, and Technology Pump-Priming Grant from Universities UK International. The co-authors would like to thank Fabrice Menneteau and Lionel Michaux (Bibliothèque Nationale de France, Département Son, Vidéo, Multimédia) for facilitating access to the collection at the heart of this paper; Anne Picaud and Claire Jahan (Institut National de Jeunes Sourds de Paris) for identifying and facilitating access to important archival materials related to this project; and Angélique Amelot (Laboratoire de Phonétique et Phonologie, UMR 7018) and Henri Chamoux (Laboratoire de Recherche Historique Rhône-Alpes, UMR 5190) for sharing their subject-specific expertise on certain aspects of the project.

7 References

- [1] The finding-aid for this collection can be found in the Bibliothèque Nationale de France’s online database ‘Archives et Manuscrits’, <https://c.bnf.fr/PQC>.
- [2] FOUGERON, C., A. AMELOT, & A. RIALLAND: *On the history of the Phonetic Institute of the Université de Paris from 1911 to 1950*. In: R. SKARNITZL & J. VOLÍN (eds.): *Proceedings of the 20th International Congress of Phonetic Sciences*, pp. 3226–3230, 2023.
- [3] SCHWEITZER, C., & C. DODANE: *De la notation musicale à la transcription de la prosodie: Description de la prosodie du français du XVI^e au début du XXI^e siècle*, *SHS Web of Conferences*, pp. 1–17, 2016.
- [4] SCHWEITZER, C.: *Des différents types d’utilisation de la notation musicale pour visualiser l’intonation du français: Entre tradition et modernité*, *Travaux de linguistique*, vol. 86, no. 1, pp. 133–156, 2023.
- [5] BERGERON, K.: *Voice lessons: French mélodie in the belle époque*, Oxford University Press, Oxford, 2010.
- [6] Bibliothèque Patrimoniale, Institut National de Jeunes Sourds de Paris, Paris.
- [7] 20010498/186, Archives Nationales, Pierrefitte-sur-Seine.
- [8] *La parole peut-être rendue à tous les sourds-muets*, *L’Ouest-Éclair*, 17 February 1914, p. 2.
- [9] PIERRE, R.: *Rendant aux muets...*, *La Petite République*, 28 February 1914, p. 4.
- [10] RIO, A.: *Les mystères de la parole dévoilés*, *Lectures Pour Tous*, 1 October 1917, pp. 754–762.
- [11] NOVERRE, M.: *Le nouvel art cinématographique*, *Le Cinéopse*, 1 October 1923, pp. 737–739.
- [12] MARICHELLE, H.: *La chronophotographie de la parole*, Atelier Typographique de l’Institut National des Sourds-Muets, Paris, 1892.

- [13] MARICHELLE, H.: *Phonétique expérimentale: La parole d'après le tracé du phonographe*, C. Delagrave, Paris, 1897.
- [14] MARICHELLE, H.: *Le mouvement et les sons de la parole*. In: A. CASTEX & R. JOUET (eds.), *Traité d'orthophonie*, pp. 11–32, J.-B. Baillière et fils, Paris, 1920.
- [15] TESTON, B.: *Le rôle d'Étienne-Jules Marey dans l'émergence de la phonétique expérimentale*. In: L.-J. BOË & C.-E. VILAIN (eds.), *Un siècle de phonétique expérimentale*, ENS Éditions, Lyon, 2010, <https://doi.org/10.4000/books.enseditions.31185>.
- [16] BROCK, G., C. BUFFONI, J. CLARENC, P.-O. GAUMIN, & F. HIRSCH: *Les laboratoires de phonétique français dans la première moitié du XX^e siècle: Instrumentations, corpus et thématiques de recherche*. In: C. DODANE & C. SCHWEITZER (eds.), *Histoire de la description de la parole: De l'instrumentation à l'instrumentation*, pp. 301–318, Honoré Champion, Paris, 2021.
- [17] *Laboratoire de la Parole à l'INJSM*, affiche, boîte M-2, Bibliothèque Patrimoniale, INJS, Paris.
- [18] DEMENÿ, G.: *Photographie de la parole*, Imprimerie Lahure, Paris, Département Science et Techniques 4-TB66-80, Bibliothèque Nationale de France, Paris, gallica.bnf.fr.
- [19] MILLS, M.: *Deaf jam: From inscription to reproduction to information*, *Social Text* 102, vol. 28, no. 1, pp. 35–58, 2010.
- [20] PISANO, G.: *Une archéologie du cinéma sonore*, CNRS Éditions, Paris, 2013.
- [21] GAILLARD, H.: *Le Laboratoire de la Parole*, *Gazette des Sourds-Muets*, 1 November 1918, p. [1].
- [22] SCHORCH, P. *Sensitive heritage: Ethnographic museums, provenance research, and the potentialities of restitutions*, *Museum & Society*, vol. 18, no. 1, pp. 1–5, 2020.
- [23] TAŃCZUK, R., & S. WIECZOREK (eds.), *Sensitive sound recordings*, *Prace Kulturoznawcze* vol. 26, no. 1, 2022.
- [24] KINSEY, A.: *Report of the proceedings of the International Congress on the Education of the Deaf*, W.H. Allen, London, 1880.
- [25] ENCREVÉ, F.: *Les sourds dans la société française au XIX^e siècle: Idée de progrès et langue des signes*, Créaphis, Grâne, 2012.
- [26] Organising Committee of the International Congress on Education of the Deaf: *A new era: Deaf participation and collaboration*, Vancouver, 2010.

A RECENT HISTORY OF THE ‘MODULATION SPECTRUM’ FOR SPEECH SCIENCE: ON THE CONTRIBUTION OF COMMUNICATION BETWEEN RESEARCH FIELDS.

Olivier Crouzet¹

¹*Laboratoire de Linguistique de Nantes, LLING, UMR6310 (Nantes Université / CNRS), France*
olivier.crouzet@univ-nantes.fr

Abstract: The *Modulation Spectrum* is a parameter of the analysis of acoustic signals that has been introduced in the field of speech perception during the 1990s. This paper addresses its origins, looking back at various research fields which have been determining in the emergence of this concept for speech communication. The primary sources seem to have emerged at the beginning of the XXth century in mathematics and engineering, providing tools to characterize and ease the transmission of signals through communication channels. Though the role of envelope information for speech perception was already investigated in the 1980s, developments of the concept of a ‘Modulation Spectrum’ occurred in parallel and may be traced back to the 1990s, with works relating to human and machine auditory perception modelling from both psychophysics / psychoacoustics and neuroscience of the auditory system, investigating the role of temporal structure representation complementing the classical spectral Fourier analysis. Though these aspects may be viewed as central in psychoacoustics and auditory neuroscience, they were only marginally taken into account in speech communication research though research on auditory perception in relation to speech have investigated these issues rather extensively. We conclude with a discussion of the importance that scientific dissemination between fields must have played in the development of this concept and discuss the conditions under which future progress in the investigation of this property for speech perception and production may be favoured.

1 Introduction: Precursors of the Modulation Spectrum

The aim of this paper is to address the description of a concept known as the ‘Modulation Spectrum’ [1, 2] (also sometimes coined the ‘Complex Modulation Spectrum’ [3] or the ‘Modulation Power Spectrum’ (MPS) [4]) in the field of speech communication research. This notion is strongly tied to the ‘Modulation Transfer Function’ (MTF) [5, 6]) and will be approached through some fundamental notions that were developed in various connected fields and gave rise to modern analytic techniques for the description of acoustic signals.

1.1 The classical Fourier decomposition

A classical description of sounds relies on Fourier spectral decomposition. Fourier analysis (1822) [7] is systematically associated with two exclusive points of view on physical signals: they may either be described in the temporal domain *or* in the spectral domain, but not both at the same time.

1. When considering a signal in the *temporal domain* (e.g. using a *waveform* representation), one will look at its pressure changes along the time dimension but will not have access to information about its frequency content,
2. When adopting a *spectral domain* viewpoint (e.g applying an FFT¹ on either the signal or a specific extract of this signal), one will focus on its *amplitude spectrum* or its *phase spectrum*, considering its frequency composition, particularly the correspondence between frequency components and their energy (the *amplitude spectrum*), but it prevents access to the time dimension.

It is well known that these two dimensions may not be accessed simultaneously. One has to turn from one dimension to the other depending on whether the temporal or the spectral information (pressure variations along the time dimension vs. amplitude of the various frequency components) is to be explored. Studying speech acoustics usually involves turning back and forth between these two representations of a sound in order to analyse their data as precisely as possible.

Though a spectrographic view of a signal will provide some simultaneous temporal and spectral information, it is constrained by what is known as the *time-frequency uncertainty principle* [8]: it is not possible to reach *both* a good temporal resolution *and* a good spectral resolution. Therefore, when one selects a wide-band spectrogram, the good timing resolution will offer the possibility to observe changes in frequency over time but spectral analysis will be limited. Alternatively, if one computes a narrow-band spectrogram, a precise spectral analysis will be provided but will partly sacrifice temporal resolution.

In any case though, time-related *changes* that would be accessible in a spectrographic representation will mostly be restricted to frequency alterations, while measurements of a change in the amplitude of the signal will hardly be accessible. Only broad variations in energy would be available from the discrimination of luminance levels in the image of the spectrogram. One way to overcome these limitations is to treat various representations of a signal simultaneously, having a concurrent look at the spectrogram(s), the spectrum and the waveform altogether.

1.2 Amplitude Modulation: engineering and mathematical sources

Before addressing the target notion of this paper (the ‘Modulation Spectrum’), we will first consider some related concepts.

It seems reasonable to relate the concept of the ‘modulation spectrum’ to the technique known as *Amplitude Modulation* of a signal. Amplitude Modulation (AM) was designed at the beginning of the XXth century as a tool for the transmission of signals over radio (wireless) communication channels. Its first use is attributed to Reginald Fessenden ca. 1899 (according to Aitken [9]) but it was further formalized by Carson [10] in 1922. ‘Amplitude Modulation’ consists in modulating the amplitude of a *carrier* (the original signal) by a signal whose amplitude varies (the *envelope*). A graphical illustration is provided in Fig. 1. This technique makes it possible to ‘demodulate’ the signal at the receiver part. Moreover, it supports the possibility to disentangle signals that would be mixed in the communication medium, provided they are modulated by different amplitude modulation frequencies, hence the possibility to transmit several radio waves from different producers and selecting only one of these on our radio receiver².

¹Fast Fourier Transform

²Note though that we most frequently use *Frequency Modulation* (FM) for listening to the radio, which is an alternate technique to *Amplitude Modulation* (AM) for radio transmission.

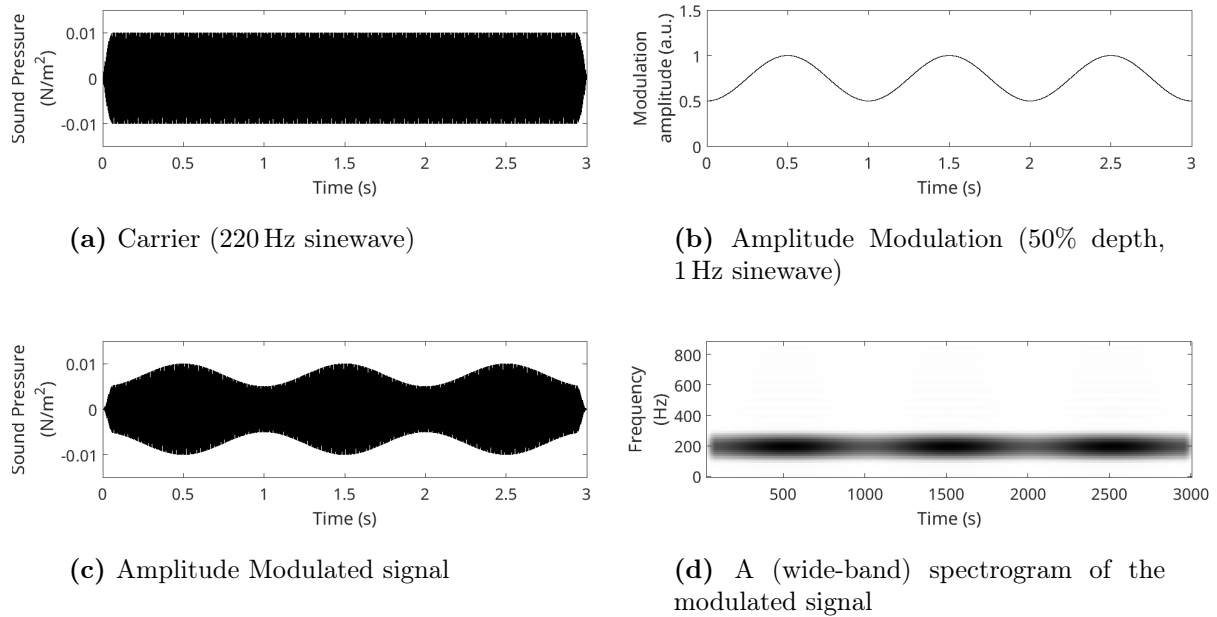


Figure 1. An example of Amplitude Modulation. Fig. 1a displays the waveform of a carrier signal (in this case a constant amplitude 220 Hz sinewave). Fig. 1b displays a view of a 1 Hz modulation sinewave with 50% depth. Fig. 1c is the result of the Amplitude Modulation applied to the carrier signal. Fig. 1d represents a wide-band spectrogram (computed on 32 FFT points) of the modulated signal. This spectrogram is aimed at maximizing temporal resolution but consequently provides very broad spectral information. One can see that it still provides only limited timing information with respect to the amplitude modulation.

One most interesting aspect of this technique considering speech communication and –more generally– natural sounds, is that this approach of a signal as the combination of a *carrier* with an *envelope* may be applied to natural signals as was conceptualized by Hilbert [11] in 1912 (cited by Smith et al. [12]). Indeed, any natural signal may be viewed as the combination of a *rapidly* varying signal with a *slowly* varying envelope. In this case the main difference with respect to what was previously described is that the envelope of the signal may perfectly be generalized to a complex wave that is directly produced by the source. The differentiation between the carrier and the envelope relates to their respective ranges of modulation velocities. The *carrier* signal is associated with oscillations that vary at relatively high frequencies and may be associated with an *auditory sensation* by the listener. The *envelope* signal is associated with slower oscillations that would not generally induce an auditory sensation for the listener but would be perceived as changes in amplitude provided they are applied on an audible sound. These will nevertheless be associated with frequencies (expressed in Hertz) and will typically be related to low frequencies. The frequency ranges characterizing these components of a natural signal are usually not strictly defined but, for the sake of illustration, one may roughly consider that frequencies below 100 Hz may be associated to the range of ‘envelope modulations’ while higher frequencies up to approximately 20kHz for the human species would be associated with the range of ‘carrier information’, though there may be some overlap.

Fig. 2 illustrates this description: The envelope of a natural speech signal was extracted using a conventional Hilbert transform using the Matlab software suite along with the Signal Processing Toolbox [13]. The original envelope was low-pass filtered with a 100 Hz cutoff frequency, therefore limiting changes in the envelope signal to modulation frequencies under 100 Hz. It is visible from the envelope information on the graphs that a diverse range of modulation velocities from lower to higher frequencies (Fig. 2c) may

characterize this signal. This fact is only partly visible on the envelope itself (Fig. 2a) which displays various paces in energy envelope increase or decrease.

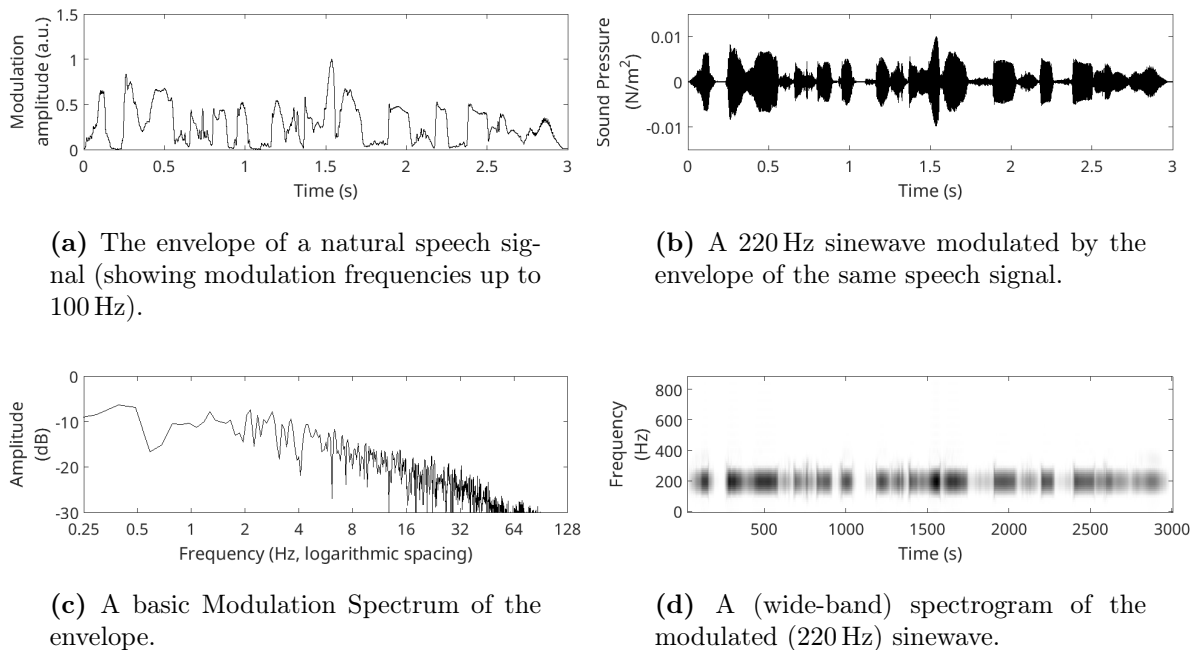


Figure 2. An example of Amplitude Modulation associated with a complex envelope modulation. Fig. 2a displays the envelope of a natural speech signal. Fig. 2b displays the result of the modulation of the original 220 Hz sinewave by this natural speech envelope. Fig. 2c represents a basic modulation spectrum of the envelope, showing that a range of low-frequencies are associated with the corresponding information. Fig. 2d represents a wide-band spectrogram (computed on 32 FFT points) of the modulated signal.

1.3 Approaches to the role of envelope information in speech communication

The process by which envelope information may provide relevant cues for speech signals was investigated by various authors during the 1980s and the 1990s (see Rosen [14]; Nguyen and Hawkins [15], for reviews). Van Tasell et al. [16] were probably among the first to show that applying the amplitude modulation of a natural speech signal on a carrier containing absolutely no information (white noise) may provide relevant cues to the identification of consonants in a 19-Alternative Forced Choice task (19-AFC). Indeed, their participants' performance was higher than random even though no spectral information was provided (up to 35% overall correct recognition in the best condition when the theoretical random performance can be estimated at $1/19 \times 100 = 5.26\%$).

Envelope-related cues were also investigated when spectral content is available. Shinn and Blumstein [17] compared conditions where a /ba/~/wa/ synthetic continuum with the appropriate spectral content was submitted to various envelope conditions. In 2 out of the 3 experimental conditions, the envelope was systematically forced to either a stop amplitude envelope or a glide amplitude one. The authors observed that the envelope had a major impact on the contrast perception, envelope cues associated with a continuant vs. non-continuant contrast ruling out the spectral cues. Similar phenomena were described by Kluender and Walsh [18] when investigating the classification of unvoiced fricatives vs. affricates. Parallel results were outlined in a sentence identification task by Smith et al. [12], though alternate interpretations were advocated, arguing that in

some specific situations, the auditory system may partly recover envelope cues from fine spectral structure information (Gilbert and Lorenzi [19]).

As a matter of fact, these investigations were limited to exploring the sole contribution of an energy envelope as applied to a *global* spectral structure (the ‘full spectrum’ of a signal –or a noise– was systematically modulated by a single complex envelope). Further explorations addressed the contributions of envelope information within different spectral channels. When a sound is decomposed into a number of different frequency bands, each of these bands is associated with an individual slowly varying complex envelope. These individual envelopes may differ at various degrees between channels and the envelope information associated with multiple channels is an efficient information provider (e.g. in cochlear implants, Shannon et al. [20]), at least for speech in the absence of noise.

However, it occurs that timing information accessible to the auditory system may be more complex than what is delineated by these studies. Indeed, we’ve previously mentioned the presence of multiple ‘modulation frequencies’ in acoustic envelopes. Therefore, talking about the role of envelope in auditory processing may be too restricted as various frequencies of this component may provide different information. Also, it is nowadays admitted that the auditory system preserves various levels of timing information even after the spectral computation that takes place in the cochlea. This is where the notion of the ‘modulation spectrum’ enters into play.

2 Toward the concept of ‘Modulation Spectrum’

During the same period of time (1920s-1990s), approaches to auditory processing in animals and humans investigated the way the auditory system processes basic sound information.

2.1 Pitch coding: place vs. time theories

One famous historical controversy that nicely relates to the issues addressed in this paper concerns how the auditory system processes *pitch* for complex sounds. Lyon [21] describes the debate that took place during 50 years from the beginning of the XXth century, involving von Helmholtz [22] and various auditory scientists (von Békézy [23, 24], Licklider [25], Johnstone et al. [26] among others). This debate relied on theoretical problems concerning the way the auditory system processes information for pitch analysis in terms of either place-coding (tonotopy) or time-coding (periodicity of stimulation). It occurred that both processes had to be implemented in models of auditory processing in order to account for the various patterns of pitch perception in complex sounds, very much as spectral analysis does not seem to be sufficient in order to account for speech perception and would require to involve the processing of various temporal information sources.

2.2 The *Auditory Image* approach: multiple sound representations in the auditory system

A new concept emerged from this controversy: the notion of ‘Auditory Image’. Partly inspired by David Marr’s research on visual processing [27], the *Auditory Image Theory* states that the representation of sounds must account for the fact that the auditory system processes acoustic information at various hierarchical levels of the brain in order to project information on 2-dimensional cortical arrays (see Lyon [21]). According to this approach, a diversity of different theoretical models may perfectly be implemented into a single larger system, each of them providing a way to solve specific issues, possibly

taking place at different locations of the auditory pathway but exchanging information with other models whose combination forms the entire auditory processing system.

In order to account for both tonotopic spectral analysis in the cochlea and the parallel importance of temporal information availability in perceptual mechanisms, the preservation of various timing cues after the cochlear spectral analysis stage has been of central importance. As stated by Lyon [21, p. 143] (talking about wide-band vs. narrow-band spectrograms):

‘But neither version captures the much finer temporal structure, the phase information or individual waveform peaks that the auditory nerve can represent.’

The importance of temporal information was soon addressed in various speech perception experiments concerning human as well as machine perception and computer modelling. This field of research introduced the notion of the ‘Modulation Spectrum’ which, according to the ‘Auditory Image’ approach principle, may be defined as a specific type of sound representation that could interact with other representations such as spectral patterns.

2.3 The Modulation Spectrum for speech perception and speech analysis

As was described in section 1.3, information in speech may combine both a spectral description (which frequencies are associated with greater magnitudes) and a temporal one (how information changes along the time dimension). Though some such changes may be visible on a spectral representation that evolves over time (e.g. formant transitions, f_0 variations...), we have seen that part of this temporal information is not available through the frequency spectrum and would involve complementary data (e.g. envelope information among others).

Fig. 2c is a very basic representation of the modulation spectrum for the envelope of a natural signal. An FFT has been applied on the envelope signal in order to identify the modulation frequencies composing this envelope. According to this visualization, it provides information concerning the various slowly-varying envelope modulations that may be viewed as quantifying temporal speeds of amplitude changes in the signal. The magnitudes in this spectrum (the vertical dimension) tell whether some temporal modulations dominate in the signal. It is noticeable that more complete representations may be computed by previously decomposing a signal into a number of frequency channels –just as the cochlea tonotopically decomposes sounds into frequencies– and then computing the modulation spectrum on each output of the filterbank. This would then offer a representation giving access to various modulation frequencies at several frequency locations in the spectrum [28, 29]. It is also possible to combine modulation spectrum over time with the characterization of modulation power over frequency, giving rise to 2-D representations that fall out of the scope of this paper (see Chi et al. [6], Elliott and Theunissen [30])

Stephen Greenberg and Hynek Hermansky contributed to this subject relatively early in speech communication research [28, 31]. Arai and Greenberg [28] collected data on human sentence identification when specific spectral ‘slits’ (frequency channels) were temporally desynchronized with one another. The authors observed that speech identification performance was correlated with the magnitude of the low-frequency range (3 to 6 Hz) of the modulation spectrum, concluding that speech intelligibility was probably strongly related to temporal modulation information. Kingsbury et al. [31] implemented temporal modulation in what they called a ‘modulation spectrogram’ which was provided as a front-end to an ASR system and measured a significant improvement in performance

on highly reverberant speech with comparison to state-of-the-art systems at the time of the experiment (see Hermansky [32] for a review). This representation was also used by Cummins and Port [33] as a tool to characterize rhythm in speech. More recently, Greenberg's seminal work was extended by Stilp et al. [29] in order to investigate the impact of speaking-rate variations on these phenomena.

3 Discussion and perspectives

The historical perspective that is described in this paper outlines the strong relationships that tied various disciplines together in order to develop the concept of the 'Modulation Spectrum': engineering, mathematics, visual psychophysics, psychoacoustics, auditory neuroscience, psychological auditory modelling, auditory computational modelling and various fields in speech sciences communicated together, either explicitly or implicitly, for this concept to emerge. It may be interesting to add to this story an outline of how communication between the people involved in this subject took place in terms of personal contacts and collaborations over both *time* and *space*.

Perspectives on the present and future of these issues are also of particular interest. According to Lyon [21, e.g. p. 94], it is still the case that temporal information is not sufficiently and adequately represented in most technical accounts of artificial audio processing (*machine hearing*). As a matter of fact, at a recent congress on the interpretability of Artificial Intelligence, I was quite surprised that all the presentations concerning speech recognition performed absolutely no auditory-based processing of their input.

It also seems to be the case that the role of the 'Modulation Spectrum' is not central in the field of speech communication, and particularly in the domain of phonetic and phonological analysis. Though a non negligible amount of research in auditory processing addresses speech perception, it is sometimes associated with partly different research questions. It may perhaps be viewed as a *niche* compared to what the spectral and global envelope cues offer in speech perception modelling. As of today however, it seems obvious that technical aspects of the computations may not be easily accessible when compared with alternative acoustic tools and representations (spectrum analysis, spectrograms, waveforms...). Progress in technical accessibility may prove crucial for these phenomena to acquire a stronger status in the future of speech communication research.

Acknowledgments

Parts of this work received funding from the program *Recherche – Formation – Innovation* 'Ouest Industries Créatives' (RFI-OIC, Région Pays de la Loire).

References

- [1] KANEDERA, N., T. ARAI, H. HERMANSKY, and M. PAVEL: *On the relative importance of various components of the modulation spectrum for automatic speech recognition*. *Speech Communication*, 28, pp. 43–55, 1999. doi:10.1016/S0167-6393(99)00002-3.
- [2] CHRISTIANSEN, T. U., T. DAU, and S. GREENBERG: *Spectro-temporal processing of speech – an information-theoretic framework*. In B. KOLLMEIER, G. KLUMP, V. HOHMANN, U. LANGEMANN, M. MAUERMANN, S. UPPENKAMP, and J. VERHEY (eds), *Hearing – From Sensory Processing to Perception.*, pp. 517–523. Springer, 2007. doi:10.1007/978-3-540-73009-5_55.

- [3] GREENBERG, S. and T. ARAI: *The relation between speech intelligibility and the complex modulation spectrum*. In *7th European Conference on Speech Communication and Technology (Eurospeech 2001)*, pp. 473–476. Aalborg, Denmark, 2001. doi:10.21437/Eurospeech.2001-125.
- [4] ARNAL, L. H., A. FLINKER, A. KLEINSCHMIDT, A.-L. GIRAUD, and D. POEPEL: *Human Screams Occupy a Privileged Niche in the Communication Soundscape*. *Current Biology*, 25(15), pp. 2051–2056, 2015. doi:10.1016/j.cub.2015.06.043.
- [5] DRULLMAN, R., J. FESTEN, and T. HOUTGAST: *Effect of temporal modulation reduction on spectral contrasts in speech*. *The Journal of the Acoustical Society of America*, 99(4, Pt. 1), pp. 2358–2364, 1996. doi:10.1121/1.415423.
- [6] CHI, T., Y. GAO, M. C. GUYTON, P. RU, and S. SHAMMA: *Spectro-Temporal Modulation Transfer Functions and Speech Intelligibility*. *The Journal of the Acoustical Society of America*, 106(5), pp. 2719–2732, 1999.
- [7] FOURIER, J.: *Théorie analytique de la chaleur*. Firmin Didot, Paris, 1822. URL <https://gallica.bnf.fr/ark:/12148/bpt6k1045508v>.
- [8] GABOR, D.: *Theory of communication*. *Journal of the Institute of Electrical Engineering*, 93, pp. 429–457, 1946. doi:10.1049/ji-1.1947.0015.
- [9] AITKEN, H. G.: *The Continuous Wave: Technology and American Radio, 1900-1932*. Princeton University Press, Princeton, N.J., 1985. URL <https://press.princeton.edu/books/hardcover/9780691639680/the-continuous-wave>. ISBN: 9780691639680.
- [10] CARSON, J.: *Notes on the Theory of Modulation*. *Proceedings of the Institute of Radio Engineers*, 10(1), pp. 57–64, 1922. doi:10.1109/JRPROC.1922.219793.
- [11] HILBERT, D.: *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*. Teubner, Leipzig, 1912. URL <https://gallica.bnf.fr/ark:/12148/bpt6k995549>.
- [12] SMITH, Z. M., B. DELGUTTE, and A. J. OXENHAM: *Chimaeric sounds reveal dichotomies in auditory perception*. *Nature*, 416(6876), pp. 87–90, 2002. doi:10.1038/416087a.
- [13] THE MATHWORKS INC.: *Matlab version: 9.13.0 (r2022b)*. 2022. URL <https://www.mathworks.com>.
- [14] ROSEN, S.: *Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects*. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 336(1278), pp. 367–373, 1992. doi:10.1098/rstb.1992.0070.
- [15] NGUYEN, N. and S. HAWKINS: *Temporal integration in the perception of speech: introduction*. *Journal of Phonetics*, 31(3), pp. 279–287, 2003. doi:10.1016/j.wocn.2003.09.003.
- [16] VAN TASSELL, D., S. SOLI, V. KIRBY, and G. WIDIN: *Speech waveform envelope cues for consonant recognition*. *The Journal of the Acoustical Society of America*, 77, pp. 1069–1077, 1987. doi:10.1121/1.395251.
- [17] SHINN, P. and S. E. BLUMSTEIN: *On the role of the amplitude envelope for the perception of [b] and [w]*. *The Journal of the Acoustical Society of America*, 75(4), pp. 1243–1252, 1984. doi:10.1121/1.390677.
- [18] KLUENDER, K. R. and M. A. WALSH: *Amplitude rise time and the perception of the voiceless affricate/fricative distinction*. *Perception & Psychophysics*, 51(4), pp. 328–333, 1992. doi:10.3758/BF03211626.
- [19] GILBERT, G. and C. LORENZI: *The ability of listeners to use recovered envelope*

- cues from speech fine structure. *The Journal of the Acoustical Society of America*, 119(4), pp. 2438–2444, 2006. doi:10.1121/1.2173522.
- [20] SHANNON, R., F. ZENG, V. KAMATH, J. WYGONSKI, and M. EKELID: *Speech recognition with primarily temporal cues*. *Science*, 270, pp. 303–304, 1995. doi:10.1126/science.270.5234.303.
- [21] LYON, R. F.: *Human and Machine Hearing: Extracting Meaning from Sound*. Cambridge University Press, Cambridge, UK, 2017. doi:10.1017/9781139051699.
- [22] VON HELMHOLTZ, H.: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Druck und Verlag von Friedrich Vieweg und Sohn, Braunschweig, 1862. URL <https://imslp.org/wiki/Special:ReverseLookup/90090>.
- [23] VON BÉKÉZY, G.: *Zur theorie des hörens: die schwingungsform der basilarmembran*. *Physik Zeits*, 29, pp. 793–810, 1928.
- [24] VON BÉKÉZY, G.: *The pattern of vibrations in the cochlea*. In G. v. BÉKÉZY and E. G. WEVER (eds), *Experiments in Hearing*, pp. 403–484. Krieger, 1960.
- [25] LICKLIDER, J. C. R.: *A duplex theory of pitch perception*. *The Journal of the Acoustical Society of America*, 23, p. 147, 1951. doi:10.1121/1.1917296.
- [26] JOHNSTONE, B. M., R. PATUZZI, and G. K. YATES: *Basilar membrane measurements and the travelling wave*. *Hearing Research*, 22(1), pp. 147–153, 1986. doi:10.1016/0378-5955(86)90090-0.
- [27] MARR, D.: *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press, 1982. doi:10.7551/mitpress/9780262514620.001.0001.
- [28] ARAI, T. and S. GREENBERG: *Speech intelligibility in the presence of cross-channel spectral asynchrony*. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 933–936. Seattle, USA, 1998. doi:10.1109/ICASSP.1998.675419.
- [29] STILP, C. E., M. KIEFTE, J. M. ALEXANDER, and K. R. KLUENDER: *Cochlea-scaled spectral entropy predicts rate-invariant intelligibility of temporally distorted sentences*. *The Journal of the Acoustical Society of America*, 128(4), pp. 2112–2126, 2010. doi:10.1121/1.3483719.
- [30] ELLIOTT, T. M. and F. E. THEUNISSEN: *The Modulation Transfer Function for Speech Intelligibility*. *PLOS Computational Biology*, 5(3), p. e1000302, 2009. doi:10.1371/journal.pcbi.1000302. Publisher: Public Library of Science.
- [31] KINGSBURY, B. E., N. MORGAN, and S. GREENBERG: *Robust speech recognition using the modulation spectrogram*. *Speech Communication*, 25(1), pp. 117–132, 1998. doi:10.1016/S0167-6393(98)00032-6.
- [32] HERMANIKY, H.: *History of modulation spectrum in ASR*. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5458–5461. IEEE, Dallas, TX, USA, 2010. doi:10.1109/icassp.2010.5494907.
- [33] CUMMINS, F. and R. F. PORT: *Rhythmic constraints on stress timing in english*. *Journal of Phonetics*, 26, pp. 145–171, 1998. doi:10.1006/jpho.1998.0070.

THE SOUND CHARTS OF FRIEDRICH RAUSCH AND THEIR CREATOR

Rüdiger Hoffmann

*Technische Universität Dresden, Institut für Akustik und Sprachkommunikation
ruediger.hoffmann@tu-dresden.de*

Abstract: With the expansion of educational opportunities for the “deaf and dumb” (today deaf-mute or simply deaf), the need for specialized teaching aids also arose. Among these, the sound charts, which the deaf and dumb teacher FRIEDRICH RAUSCH began publishing in 1903, stand out. This paper describes the publishing history of the sound charts along with RAUSCH’s life as a teacher. Later, RAUSCH ran a successful company producing various other teaching aids, which will be discussed only briefly.

1 Introduction

This article is part of the ongoing investigation of the exhibits in the Historical Acoustic-Phonetic Collection (HAPS) at TU Dresden. A significant portion of the objects from the history of phonetics deal with the representation and study of human articulation. The development of suitable teaching materials was driven primarily by deaf education, but also by vocal pedagogy. Impressive three-dimensional, plastic teaching aids were created, such as the articulation model by the singing teacher AUGUSTE BÖHME-KÖHLER (1856–past 1936) since approx. 1895 (Fig. 1), which was also available in a particularly large version for the classroom. For use in teaching, however, graphic representations of the articulation of various sounds were also required which could be found in various textbooks (e. g., Fig. 2). Additionally, large format versions were needed for the classroom. An early example emerged at the “Institute for the Deaf and Mute” in Munich, where 16 large-format “sound charts” were created based on designs by director MAX KOLLER and executed by the artist GUSTAV KELLER (1860–1911). He was a well-known painter and illustrator of scientific works. Unfortunately, it is unknown when these charts were manufactured. They are mentioned in a report from 1906 [10, p. 17] and printed as small-sized black-and-white reproductions in an exercise book from 1908 (Fig. 3) [21].

The invention of X-ray technology in 1895 led to more precise knowledge of articulation, particularly through the work of the physician MAX SCHEIER (1864–1921) [22]. During this period, in 1903, the deaf-mute teacher FRIEDRICH RAUSCH (1871-1956) from the town Nordhausen in Thuringia came up with the idea of producing large-format sound charts for school use as a publishing product in large quantities [18]. Because a complete set from the 2nd edition (1911) has been preserved in the HAPS, initial information about F. RAUSCH was collected for the list of namesakes [12, p. 172], and the research continued thereafter. This study was largely completed by the 2023 Phonetics Congress in Prague, but due to limited space, it could only be briefly mentioned in the author’s contribution [8].

This investigation took on the character of an adventure tour through several archives. It emerged that F. RAUSCH was known as a producer of teaching materials and, in this context, was valued as a representative of “experimental archaeology” [23]. The origin of the sound charts, however, remained obscure. The known biographical information was based on a type-written CV, which proved to be selective and therefore unreliable [27]. A breakthrough came when F. RAUSCH’s personnel file as a teacher was discovered in the Gotha State Archives [28]. With its help, the biography could be reconstructed. The results were published in German in a detailed article in a yearbook of the city of Nordhausen [9]. The present article is a 50 percent



Figure 1. (left): Model for sound formation from A. BÖHME-KÖHLER, left cheek removed. Hamburg Phonetics Laboratory, now in TU Dresden / HAPS. Photograph by R. DIETZEL, reproduced from the HAPS catalogue [11, p. 187].

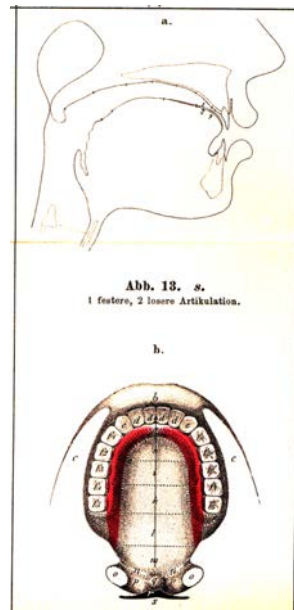


Figure 2. (middle): One of the graphics (sound [s]) in the book on German phonetics by OTTO BREMER 1893 [2].

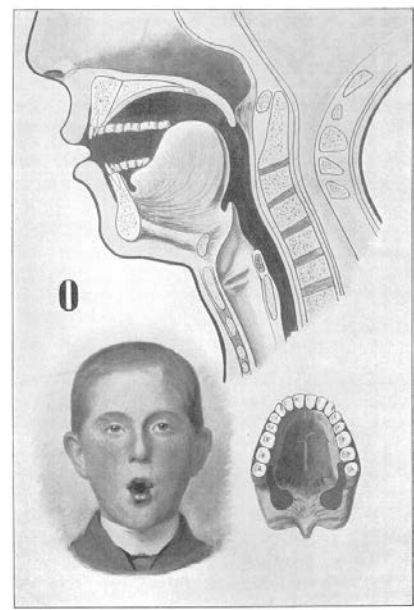


Figure 3. (right): One of the sound charts (sound [o]) of the Munich Institute for Deaf and Mute, reproduced from [21].

shortened English version, which summarizes the life story of F. RAUSCH (Sections 2 and 4) and the editorial history of the sound tables (Section 3). Detailed information on the biography with extensive references can be found in the original article [9], while F. RAUSCH's activities as a producer of teaching materials have been described not only in [23], but also later in [4].

2 The Way of Friedrich Rausch as a Teacher

2.1 Education and Professional Years in Halle

FRIEDRICH RAUSCH (Fig. 5) was born on April 21, 1871, in Treffurt (Fig. 4), a small town in the Prussian province of Saxony, where almost his entire life would be spent (Fig. 6). His father was a farmer; church records refer to him as a smallholder or ploughman. FRIEDRICH was the youngest of 11 children (two of whom died as infants) [26]. He was destined for a career as an elementary school teacher.

FRIEDRICH RAUSCH began his training at the Royal Protestant Elementary School Teachers' Seminary in Erfurt. The period of training at seminaries was three years in Prussia, usually preceded by two years of preparatory school [14]. Whether the latter applied to RAUSCH is unknown; on August 31, 1892, he received the "certificate of qualification as an elementary school teacher for provisional employment". He had to take the second examination within five years (and did it on November 3, 1894).

He made double use of his education time: In the last year of the seminary, he was a student at the deaf-mute institute in Erfurt and received a "certificate of qualification for teaching the deaf and dumb" on August 12, 1892.

On October 1, 1892, FRIEDRICH RAUSCH took up a position as a second teacher at the Dietzhausen elementary school (near the town Suhl). From this position, he attempted to utilize his qualifications as a teacher of the deaf and dumb, and on November 21, 1893, he



Figure 4. (left): The former RAUSCH family estate in Treffurt, Ziddelstraße 20. Photo: Treffurt Library / SABINE ROTH, 2024.

Figure 5. (right): FRIEDRICH RAUSCH in his later years. Photograph in the Treffurt Local History Museum.

submitted a request for employment at an institution for the deaf and dumb to Count VON WINTZINGERODE, the governor of the Prussian Province of Saxony at Merseburg. On March 3, 1894, the governor responded with a letter to the Royal Government in Erfurt, citing a high number of applications from deaf and dumb pupils to the Provincial Institute for the Deaf and Dumb in Halle. The Department of Church and School Affairs subsequently dismissed RAUSCH from his position in Dietzhausen effective March 31, and he received his desired position as an assistant teacher in Halle on April 1, with a starting salary of 1,200 marks.

Regarding RAUSCH's work at the Institute for the Deaf and Dumb, his graduation certificate states that he "distinguished himself during this time both through his diligent pursuit of good performance in the pupils department assigned to him, and through his conscientious fulfillment of his other official duties". During his time there, he worked on his qualifications. The curriculum vitae states [27]: "[...] in his spare time, he studied at the University of Halle."

It is tempting to consider this brief remark as the key to the qualifications that the later FRIEDRICH RAUSCH demonstrated beyond doubt in both the fields of speech therapy/phonetics and cultural history. Unfortunately, only one thing can be proven with certainty: He enrolled as a guest student in philosophy at the University of Halle on April 20, 1894, for the summer semester of 1894, and only for that [29].

Therefore, it is pure speculation what RAUSCH might have studied alongside his work during the relatively short period in question, 1894–1896. One hypothesis arises from the fact that he must have somehow acquired the phonetic background for his sound charts: The speech scientist OTTO BREMER (1862–1936), already mentioned in Fig. 2, worked as a private lecturer at the German Seminar since 1888 and may well have developed ideas for the sound charts and passed them on to RAUSCH.

A positive assessment of his educational achievements can be found in the certificate issued by the director of the institution on November 20, 1896, although it does not mention the university [28]: "At the same time, he has always made a sincere effort not only to acquire the theoretical knowledge required for teaching the deaf and dumb, but also to delve deeper into pedagogy in general. His skill in producing scientific specimens deserves special attention." This corresponds precisely to the foundation for his later work in Nordhausen.

The reason for this testimony is bitter, albeit self-inflicted: On October 27, 1896, the father of a pupil at the school for the deaf and dumb complained to Governor Count VON WINTZINGERODE about the repeated corporal punishment his son had been subjected to by

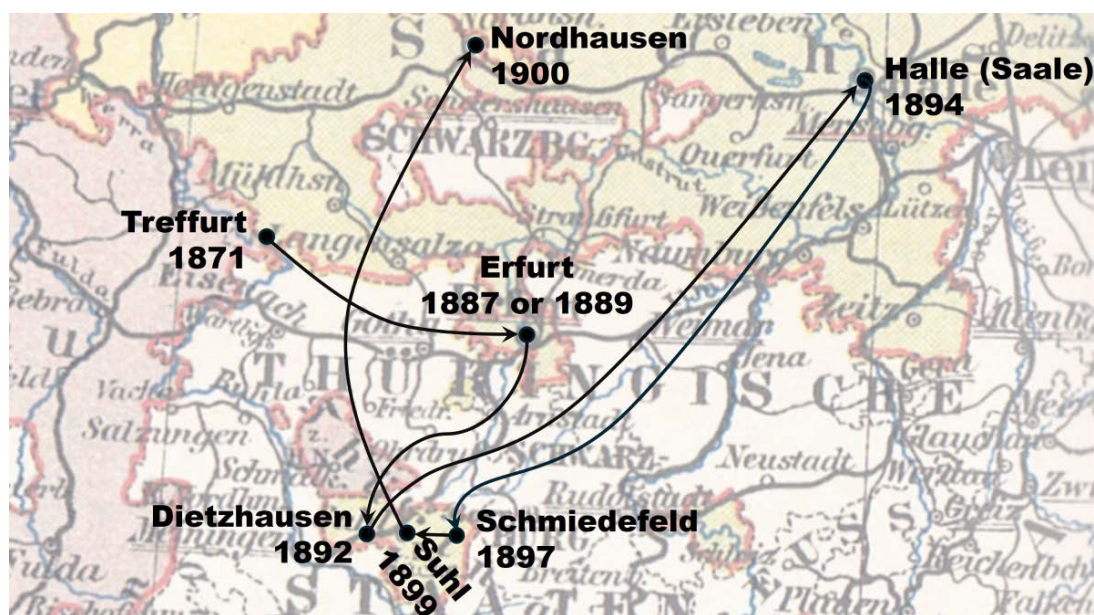


Figure 6. Parts of the Prussian Province of Saxony (administrative districts Erfurt and Halle) with the stations of life of FRIEDRICH RAUSCH. Background map from [14].

teacher RAUSCH. After extensive written back and forth, RAUSCH received a letter of termination on November 14, 1896, for exceeding the law on corporal punishment. The notice period was six months, meaning that he needed a new job by May 15, 1897, at the latest.

2.2 From the Career Setback to a Teaching Position in Nordhausen

Even before his notice period expired, FRIEDRICH RAUSCH found a position at the Protestant elementary school in Schmiedefeld on April 1, 1897. On October 1, 1899, after requesting a transfer, he began working at the public school in Suhl. At the same time, he applied for positions in more attractive locations, such as Stettin and Magdeburg. While these applications were unsuccessful, he received a position at the elementary school in Nordhausen on October 1, 1900. Nordhausen was a former “imperial city” with nearly 30,000 inhabitants and a significant industry. Despite RAUSCH started numerous additional activities in Nordhausen, he retained his teaching position, presumably because it formed his financial backbone. This was especially true during the First World War, as he reported to his publisher [25]. Only at the age of 54 did the double burden end due to health reasons. His request for retirement was granted, effective October 1, 1924.

RAUSCH must have begun realizing his ambitions parallel to his arrival in Nordhausen, because the first edition of his sound charts appeared as early as 1902, marking the resumption of his work for the deaf and dumb (cf. Sec. 3). He also published two papers in the leading journal on speech therapy [16] [17]. Shortly later, in 1905, he founded a publishing house for teaching materials (cf. Sec. 4).

We have only sketchy information about the practical aspects of RAUSCH’s work. One project about which we can provide some more details is the establishment of a kindergarten for the deaf and dumb in Nordhausen, which was announced in an information leaflet and accompanied by a request for support [31]. RAUSCH claims to have realized the idea with his sister MATHILDE (born 1864 [26]), a certified kindergarten teacher and trained deaconess with experience in treating deaf and dumb people. After 1905, there is no evidence that the kindergarten project was continued, how successful it may have ever been.

RAUSCH subsequently changed his approach and established a curative education institution for the hard of hearing and speech-impaired. It is listed in the address books from

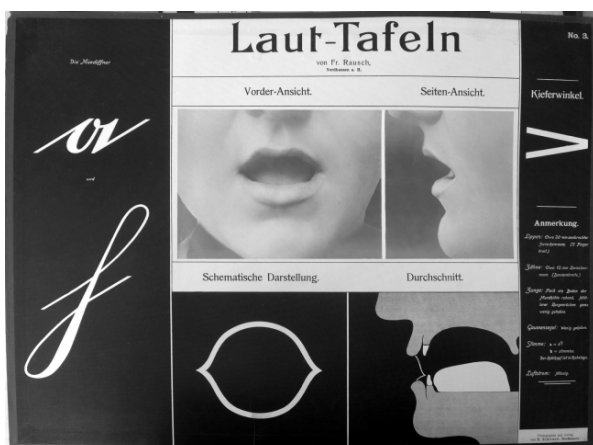


Figure 7. (left): Example of a sound chart from the first, landscape-format edition, Nordhausen 1903 (Chart 3, sounds [a, h]). Reproduction from the only known copy in the ULB Sachsen-Anhalt, Halle/Saale.

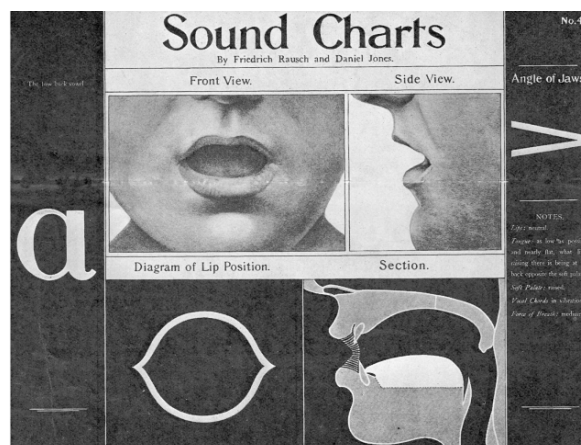


Figure 8. (right): Example of a sound chart from the English edition, London 1911 (Chart 4, sound [a]). Cutout from an advertising flyer from the publisher J. M. Dent & Sons. Flyer in the offprint collection of G. PANCONCELLI-CALZIA, now TU Dresden, HAPS.

1906/1907 to 1911/1912. It is notable that the term “deaf-mute” is no longer used, thus addressing a larger target group. By the time of World War I at the latest, RAUSCH will have abandoned his curative education activities in favor of his teaching materials producing company.

3 Edition History of the Sound Charts

3.1 The First Edition in Nordhausen

With his sure instinct for pedagogical tools, FRIEDRICH RAUSCH must have known that his sound boards were aimed at a broad audience, not just for the deaf and dumb. He subtitled the accompanying brochure [18]: “For use in reading and writing instruction based on phonetic principles, in singing lessons, in elocution lessons, in articulation instruction for the deaf and dumb, in speech and reading instruction for the deafened, in speech therapy instruction for those with speech disorders, and in instruction in special needs schools.” The development of the boards required thorough scientific and technical preparation.

First, RAUSCH had to decide on a system of speech sounds, which the sequence of the tables would follow. The sound system had been thoroughly researched towards the end of the 19th century, with the Marburg modernist WILHELM VIËTOR (1850–1918) making a particular contribution to its systematization. According to RAUSCH, he was particularly inspired by the textbooks of the Leipzig lecturer of general linguistics, FRIEDRICH TECHMER (1843–1891) [24], and HUGO HOFFMANN (1857–?), a teacher of the deaf and dumb from Ratibor [6]. Based on this, RAUSCH decided on a sequence of 20 tables. The phonetic peculiarities of some English and French sounds were taken into account.

Second, it was necessary to document the appearance of the (lower) facial area during the articulation of the corresponding sound. Unlike the example shown in Fig. 3, RAUSCH opted to take a front and side view, supplemented by a sketch of the mouth position. This led to a collaboration with the photographer CARL SCHIEWEK (1871–1934), who ran a photographic studio in Nordhausen under the name of his father, ROBERT SCHIEWEK.

Thirdly, the didactically very important “organ cross-sections” must be addressed. The aim was to create designs that were scientifically precise, yet clear enough to be easily understood by students. Therefore, it was not possible to simply adopt illustrations from phonetic literature,



Figure 9. (left): German education exhibition, entrance situation (“Courtyard of honour”), at the Louisiana Purchase Exposition. Reproduced from [1].



Figure 10. (right): The golden medal of the Louisiana Purchase Exposition. Reproduced from [1].

such as those found for the vowels in another work by the aforementioned HUGO HOFFMANN [7]. It must be acknowledged that RAUSCH created graphics of great memorability.

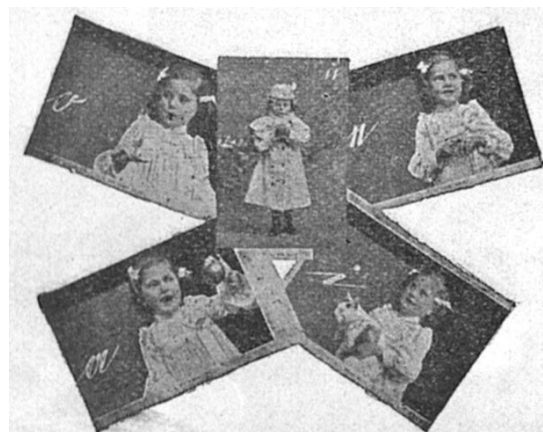
As already noted in the introduction, the possibilities for examining the internal articulatory organs had improved significantly with the discovery of X-rays. RAUSCH notes in the accompanying booklet: “The cross-sectional images are based on the X-ray photographic observations that Dr. SCHEIER made of the speech organs.” However, in the year under review, 1902, no published X-ray images of the speech process were available. This was for technical reasons; the films of that time would have required an exposure time that a subject could endure while remaining motionless, but not while articulating. Instead, the X-ray image was observed on a fluorescent screen. SCHEIER published a major (unillustrated) paper in 1898 [22], which was certainly available to RAUSCH. It described many new observations, particularly regarding the role of the soft palate. The publication of corresponding X-ray photographs began a few years later, following the technological progress.

The twenty sound charts created on this basis bear the year 1902, while the accompanying booklet is dated 1903, which also corresponds to the contemporary bibliographical records. The format of the charts, mounted on cardboard, is 48 cm × 64 cm in landscape folio. Since the subsequent edition, described below, nominally begins with the third thousand, a maximum of 2,000 copies could have been produced in Nordhausen, although the actual number remains entirely unclear. The search for specimen copies is difficult for all editions. In some cases, only the accompanying booklets are preserved. Only one publicly accessible copy of the first edition appears to exist, in the State and University Library Halle (Fig. 7).

The sound charts were shown at the 1904 World Exhibition in St. Louis. The exhibition, officially called the *Louisiana Purchase Exposition*, was of unprecedented scale. The equally extensive participation of German exhibitors meant that smaller exhibits, such as printed matter, were not explicitly listed in the official catalog. The sound charts were part of the joint exhibition “German Education” (Fig. 9).

Participation was crowned by the award of a Gold Medal (Fig. 10). One can critically note that medals were awarded in excessive numbers; the report states [1]: “The results of the award distribution were unusually favorable for Germany. Aside from the art department,

Figure 11. Advertising image for RAUSCH's "Sound Postcards". Reproduced from the third cover page of the accompanying booklet to the 2nd edition of the sound charts, Marburg 1913.



no fewer than 93 percent of all German exhibitors received awards [...] 1,577 awards were presented to German exhibitors, [...] including] 603 Gold Medals.” On the other hand, the report documents a complex judging process by an international jury, so it would be unfair to downplay the achievement. The medal was awarded to “Teacher FRIEDRICH RAUSCH and Publisher SCHIEWEK, Nordhausen.” From then on, both institutions proudly included an image of the medal on their business papers.

During this period, further works by RAUSCH appeared that are known only from bibliographies. First, there is “2 Speech Charts for Differentiating the 3rd and 4th Cases”, published by K. G. Th. Scheffer in Leipzig in 1906. Also only indirectly verifiable are the “Sound Postcards”, which show children articulating a sound “as interjection”. Their creation time is unknown, but before 1911 (Fig. 11). Their pedagogical value was likely rather limited; a first series (sounds u, o, a, e, i) and a second series (sounds ü, ö, ä, m, n) were published.

Finally, RAUSCH successfully endeavored to expand his table work to include the English language. Although details are unknown, a collaboration was established with the later important British phonetician DANIEL JONES (1881–1967), who had been working at University College London since 1907. Nine phonetic tables of English vowels by F. RAUSCH and D. JONES were published in 1911 by J. M. Dent & Sons, London, in Dent’s Modern Language Series. A promotional flyer from the publisher featuring one of the tables is known (Fig. 8). According to this publication, the tables are designed in landscape folio format (18 × 23 inches), like the original German edition. The catalog of the British Library, London, lists a specimen copy; the tables are also mentioned in JONES’s biography [3].

3.2 The Large Editions at the Publisher N. G. Elwert in Marburg

The visibility of the sound charts was enhanced by the transition to an established academic publisher, the N. G. Elwert company in Marburg. The *N. G. Elwert’sche Verlagsbuchhandlung Marburg* originated from a university bookstore founded in 1783 and had been owned by the BRAUN family since 1878. Its profile was significantly shaped by the needs of the University of Marburg, so that numerous linguistic works were also published.

The second edition of the sound charts was published by Elwert in 1911 as the 3rd to 12th thousand. The most striking innovation is the switch to a folio portrait format of 48.5 cm × 61 cm (Figs. 12 and 13). In terms of content, as stated in the accompanying booklet, “a frequently expressed wish [...] was taken into account by removing the foreign language sounds from the charts used for German and special education classes” [19]. Otherwise, the arrangement of the twenty charts remained unchanged.

The edition immediately received prominent support from GIULIO PANCONCELLI-CALZIA (1878–1966). He had worked with WILHELM VIËTOR in Marburg since 1906 and headed the phonetic laboratory of the German Colonial Institute in Hamburg since 1910. In 1911, he wrote

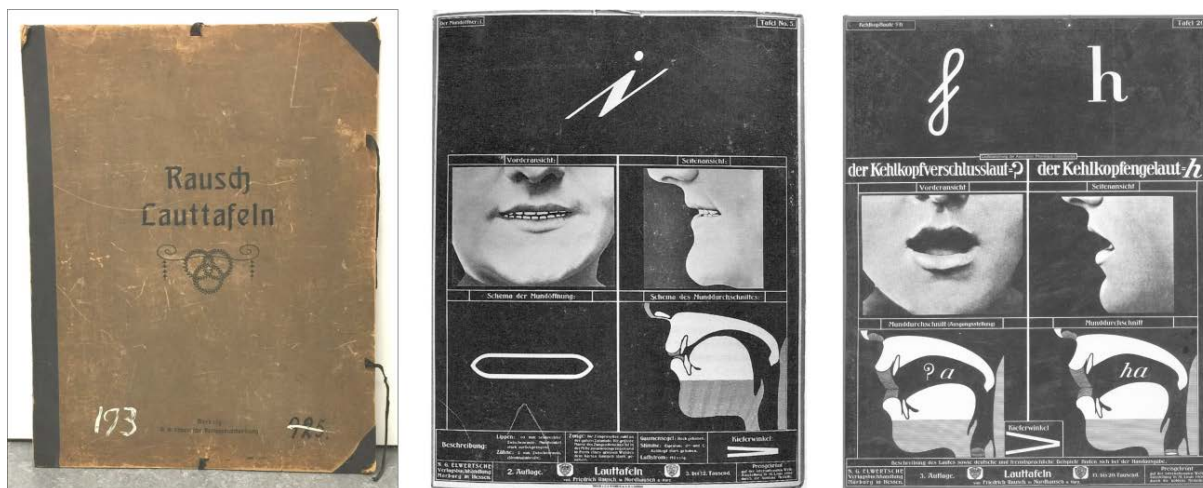


Figure 12. (left): Storage folder for the charts of the second, portrait-format edition. TU Dresden, HAPS [11, p. 204].

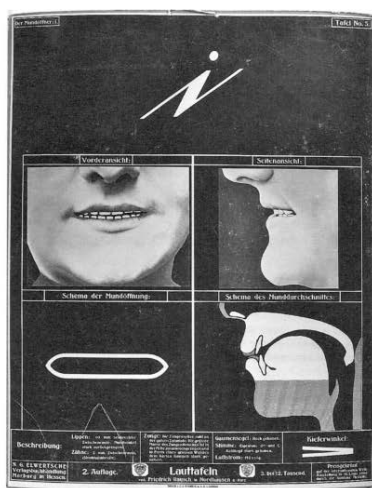


Figure 13. (middle): Example of a sound chart from the second edition, Marburg 1911 (Chart 5, sound [i]). Reproduction from the only known copy from the Hamburg Phonetics Laboratory, now TU Dresden, HAPS.



Figure 14. (right): Example of a sound chart from the third edition, Marburg 1926 (Chart 20, additionally included, sound [h] and glottal stop [ʔ]). Reproduction from the only known copy in the Schweizerisches Sozialarchiv, Zürich [30], Signature F 5153-Oa-036.

a highly positive review of the sound charts, stating that they “stand out among previous sound tables for their meticulous execution, clarity, and accuracy” [15].

The previous, rather fragmentary references to English and French sounds have been replaced by a supplementary offering. This Part II consists of six charts (3 French, 3 English), so that the complete work now consists of 26 tables.

Given the print run of 10,000 copies, it is initially surprising that no copies can be found in public libraries. The large-format plates were not intended for libraries, but for classrooms, where they were worn out. A complete set of 26 plates, originally part of PANCONCELLI-CALZIA’s laboratory equipment, arrived with that in 2006 at the HAPS of TU Dresden.

A file of the publisher in Marburg [25] reveals that RAUSCH planned various other innovations, including a Russian edition of the sound charts. The First World War thwarted these plans. In the 1920s, when RAUSCH’s work had long since shifted to his cultural-historical teaching models, the sound charts were revised for the last time. In 1926, N. G. Elwert published the third edition of the large-format tables, the 13th to 20th thousand. The only known set of these tables is in the Swiss Social Archives [30]. The Library for the Hearing and Speech Impaired in Leipzig possesses the accompanying booklet [20], which explains that the [h] sound, which had previously been treated together with the [a] sound, received a separate table together with the glottal stop (Fig. 14), so that the complete work now consists of 27 tables. Additionally, the new tables included the phonetic transcriptions of the sounds, thus tempering problems arising from the pure orthographic notation in Sütterlin or Latin script, resp. However, it must be remembered that this work was carried out before MENZERATH and LACERDA drew attention to issues of coarticulation in 1933.

This concludes the history of the edition of FRIEDRICH RAUSCH’s sound charts. Numerous reviews of the several editions of the sound charts are collected and cited in the longer German version of this paper [9]. The current owner of *Elwert Antiquariat und Verlag* in Marburg recalled in conversation that the publisher still sold remaining copies of the tables in the 1950s and 1960s [32].

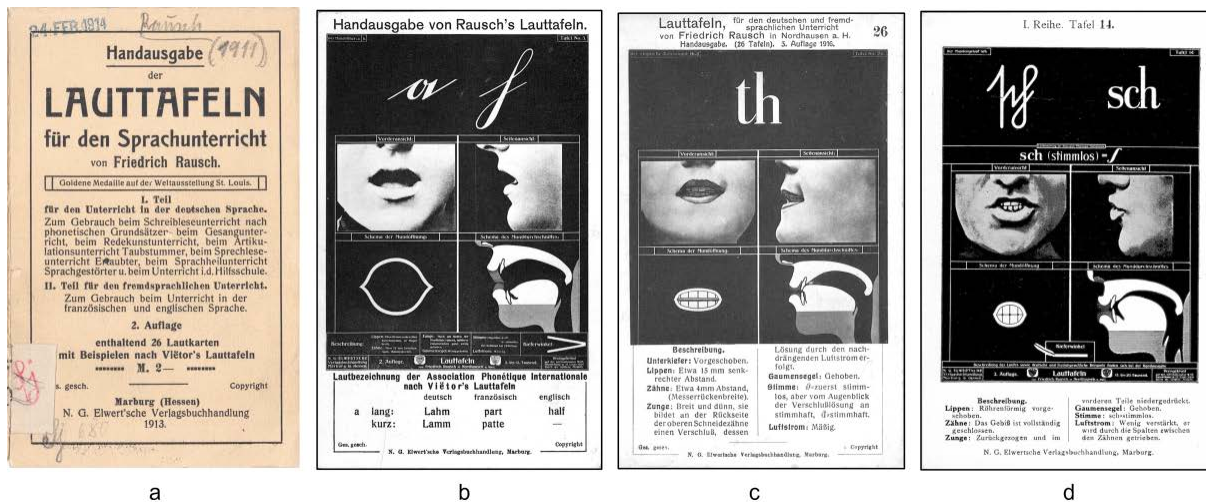


Figure 15. Examples from the known issues of the postcard-sized pocket editions. (a) Envelope from the 2nd edition, 1913; (b) Card 3 (sounds [a, h]) from the 2nd edition, 1913; (c) Card 26 (English sound [θ]) from the 3rd edition, 1916; (d) Card 20 (sound [ʃ]) from the 4th edition, 1927.

3.3 The Pocket Editions at the Publisher N. G. Elwert in Marburg

A very successful idea was to publish a “pocket edition” alongside the large-format edition, which contained the sound charts in postcard format in a folder and was available for a low price. Four editions have been published, and the most important information for each is compiled here:

1911 **1st edition:** Although no example of this edition is physically known, bibliographic evidence shows that the edition appeared in 1911 and included all 26 plates of the simultaneous “large” edition, i.e., the German and the foreign-language sections. The review by PANCONCELLI-CALZIA mentions that the cards additionally include “phonetic transcriptions” based on VIËTOR’s sound tables [15].

1913 **2nd edition:** This edition is documented by a copy in the estate of PANCONCELLI-CALZIA which is now in the HAPS (Figs. 15a and 15b). It can be seen that the original sound charts were reproduced, and a box with sample words was added underneath. The backs were used for advertising by the Elwert publishing house.

1916 **3rd edition:** During the war, this improved edition of the pocket edition was published. A complete copy of this edition is located in the Library for the Hearing and Speech Impaired in Leipzig, while the HAPS merely has ten individual cards acquired from antiquarians (Fig. 15c). In this version, the verbal descriptions of details of sound formation, which formerly were barely legible due to their small size, were reproduced in normal font at the bottom of the card. The example words previously placed there were moved to the back of the card.

1927 **4th edition:** This edition differs from the previous one in that it uses the templates of the 3rd “large” edition (Fig. 15d), thus also including a separate card for the sound [h] and the glottal stop [ʔ]. Complete copies are available at the Library for the Hearing and Speech Impaired in Leipzig and at the HAPS, for which a piece could be recently acquired from an antiquarian bookseller.

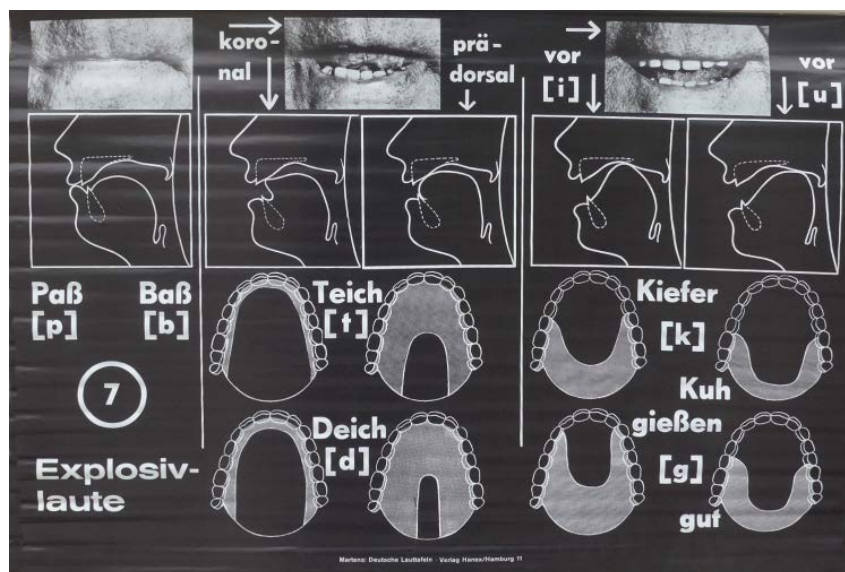


Figure 16. Example of the sound charts of MARTENS. Here: Plosive sounds. TU Dresden, HAPS.

3.4 The Later Development of Sound Charts

The concept of sound charts survived a little longer. After 1961, the Hamburg educators and phoneticians CARL and PETER MARTENS published 31 scrolls measuring 83 cm × 58 cm with the Hanex publishing house in Hamburg, which were intended to serve similar purposes as RAUSCH's charts. However, they take into account far more phonetic details (Fig. 16). A complete set is kept at the HAPS [11, p. 206]. They likely document the end of a specific type of teaching aid that would soon be replaced by new audiovisual media.

4 The Later Years of Friedrich Rausch

RAUSCH founded his teaching materials publishing company in 1905. His offerings soon expanded beyond the average school supply and included a wide variety of illustrations, models, and replicas from the entire history of culture. He authored a nine-volume "Descriptive Cultural History", intended for teachers and published between 1927 and 1931. Parallel to this, he published illustrated catalogs of teaching materials (Fig. 17), which provide a useful overview today, since few original objects have survived.

The scientific standard of his work was so high that the archaeologist GUNTER SCHÖBEL considered him one of the pioneers of experimental archaeology [23]. He also equipped and supervised several local history museums.

The company had up to 60 employees at times, but also experienced difficult economic periods. At the outbreak of the First World War, the workers had to be laid off [25]. After the beginning of the Second World War, the company switched to the production of sanitary ware. During the bombing of Nordhausen in 1945, the company was completely destroyed, so no archives are available. The Nordhausen historian HANS-JÜRGEN GRÖNKE has compiled the remaining traces [4].

After the war, the company was re-established and continued by family members to provide educational institutions with "ordinary" school supplies. The company existed in a modified form until 1973. FRIEDRICH RAUSCH himself continued to be active in a variety of roles, including for the Cultural Association and the Thüringer Landeszeitung newspaper, and also as director of the local history museum in Jüterbog. He died on May 30, 1956.

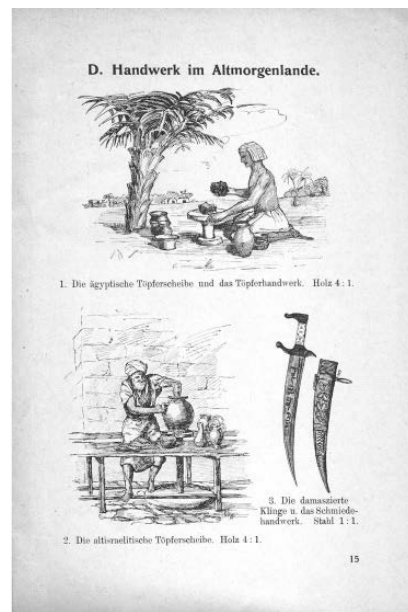


Figure 17. Example of the brochures describing the teaching objects produced by RAUSCH's company. Here: Peoples of the ancient Orient.

5 Conclusion

The publicist THEODOR HEUSS (1884–1963), who became the first President of the Federal Republic of Germany in 1949, collected some of his biographical sketches (including one about WOLFGANG VON KEMPELEN) in a book in 1948 with the subtitle "Marginal Figures of History". There, he describes the interest in "grasping historical figures that move on the margins of historical events, perhaps even intervening in major events here and there, but are characterized more by their time than they are important for it" [5, p. 10]. With regard to the history of speech research, FRIEDRICH RAUSCH is in good company here.

References

- [1] *Amtlicher Bericht über die Weltausstellung in St. Louis 1904*. Erstattet vom Reichskommissar. Berlin: Reichsdruckerei 1906 (Verhandlungen des Reichstages; 225).
- [2] BREMER, O.: *Deutsche Phonetik*. Leipzig: Breitkopf & Härtel 1893 (Sammlung kurzer Grammatiken deutscher Mundarten; 1).
- [3] COLLINS, B.; MEES, I. M.: *The Real Professor Higgins. The Life and Career of Daniel Jones*. Berlin / New York: Mouton de Gruyter 1998.
- [4] GRÖNKE, H.-J.: Eine Spurensuche in Nordhausen. Der Lehrmittelverlag Friedrich Rausch im 20. Jahrhundert. *Beiträge zur Geschichte aus Stadt und Landkreis Nordhausen, vol. 47*. Nordhausen-Salza: Verlag Steffen Iffland 2022, pp. 123–134.
- [5] HEUSS, T.: *Schattenbeschwörung. Randfiguren der Geschichte*. Tübingen: Wunderlich, 3rd edition 1960.
- [6] HOFFMANN, H.: *Einführung in die Phonetik und Orthoepie der deutschen Sprache*. Marburg: N. G. Elwert 1888.
- [7] HOFFMANN, H.: *Die Lautwissenschaft (Phonetik) und ihre Verwendung beim muttersprachlichen Unterricht in der Schule*. Breslau: F. Hirt 1901.
- [8] HOFFMANN, R.: Early activities in the field of X-ray imaging in experimental phonetics and speech training. *Proceedings of the 20th International Congress of Phonetic Sciences, Prague, August 7-11, 2023*, pp. 2219–2223.
- [9] HOFFMANN, R.: Der Taubstummenlehrer Friedrich Rausch (1871–1956) und seine Lauttafeln. *Beiträge zur Geschichte aus Stadt und Landkreis Nordhausen, vol. 49*. Nordhausen-Salza: Verlag Steffen Iffland 2024, pp. 88–108.

- [10] Königliches Zentral-Taubstumm-Institut: *Jahresbericht 1905/06*. München 1906.
- [11] MEHNERT, D.: *Historische phonetische Geräte. Katalog der historischen akustisch-phonetischen Sammlung (HAPS) der Technischen Universität Dresden, erster Teil*. Dresden: TUDpress 2012 (Studientexte zur Sprachkommunikation; 62).
- [12] MEHNERT, D.; PÉTURSSON, M.; HOFFMANN, R.: *Experimentalphonetik in Europa. Institute, Wissenschaftler, Leistungen Ende des 19. / Anfang des 20. Jahrhunderts*. Dresden: TUDpress, 2016 (Studientexte zur Sprachkommunikation; 85).
- [13] MENZERATH, P.; LACERDA, A. DE: *Koartikulation, Steuerung und Lautabgrenzung. Eine experimentelle Untersuchung*. Berlin/Bonn: Dümmler 1933 (Phonetische Studien; 1).
- [14] *Meyers Konversations-Lexikon. Eine Enzyklopädie des allgemeinen Wissens*. 4th edition in 16 volumes. Leipzig / Wien. Bibliographisches Institut 1890.
- [15] PANCONCELLI-CALZIA, G.: Annotationes phoneticae, 1911 No. 3 and No. 8/9. *Medizinisch-pädagogische Monatsschrift für die gesammte Sprachheilkunde* 21 (1911) 3, pp. 80–82, and 21 (1911) 8/9, pp. 291–294.
- [16] RAUSCH, F.: Sprechlesen. *Medizinisch-pädagogische Monatsschrift für die gesammte Sprachheilkunde* 12 (1902) 5, pp. 147–150.
- [17] RAUSCH, F.: Wert der Belehrungen über Mimik beim Sprechleseunterricht. *Medizinisch-pädagogische Monatsschrift für die gesammte Sprachheilkunde* 13 (1903) 3/4, pp. 73–78.
- [18] RAUSCH, F.: *Laut-Tafeln zum Gebrauch beim deutschen und fremdsprachlichen Sprech-Schreib-Leseunterricht nach phonetischen Grundsätzen*. Nordhausen: Schiewek 1903.
- [19] RAUSCH, F.: *Lauttafeln für den Sprachunterricht*. 2. Auflage (3. bis 12. Tausend), Marburg: N. G. Elwert 1911.
- [20] RAUSCH, F.: *Lauttafeln für den Sprachunterricht*. 3. Auflage (13. bis 20. Tausend), Marburg: N. G. Elwert 1926.
- [21] RÖTZER, F. X.: *Übungsbuch für Schwerhörige und Ertaubte*. München/Berlin: Oldenbourg 1908.
- [22] SCHEIER, M.: Über die Bedeutung der Röntgenstrahlen für die Physiologie der Sprache und Stimme. *Phonetische Studien* 11 (1897/98) = *Die neueren Sprachen* 5, Supplement, pp. 40–56.
- [23] SCHÖBEL, G.: Nachbildungen archäologischer Funde als Lehrmittel für Museen, Universitäten und den Schulunterricht in Deutschland in der Weimarer Zeit (1918–1933). *Experimentelle Archäologie in Europa, Jahrbuch 2021*. Unteruhldingen 2021, pp. 143–167.
- [24] TECHMER, F.: *Phonetik. Zur vergleichenden Physiologie der Stimme und Sprache*. Textband und Atlasband. Leipzig: Engelmann 1880.
- [25] LA Hessen, StA Marburg, Bestand: 334/2, Nr. 286 (Herstell. u. Vertrieb der “Laut-Tafeln für Sprech-, Schreib- und Leseunterricht” von F. Rausch).
- [26] LKirchenA Magdeburg, Rep. R03, Kirchenbuchfilme aus der Kirchenprovinz Sachsen, Filme 06726 (Kasualienregister Treffurt 1852-1864) u. 06727 (Kasualienregister Treffurt 1865-1877).
- [27] LA Thüringen, HStA Weimar, Bestand: Archivberatungsstelle der Provinz Sachsen, Nr. 7.
- [28] LA Thüringen, StA Gotha, Bestand Regierung Erfurt in Thüringen, Nr. 29627: Personalakten des ev. Lehrers Friedrich Rausch in Nordhausen. – Enthält auch die Personalakte der Landesdirektion der Provinz Sachsen für die Zeit als Taubstummlehrer in Halle (gesonderte Paginierung).
- [29] UA MLU Halle-Wittenberg, Rep. 4 (Rektorat der Universität Halle-Wittenberg von 1817 bis 1948), Nr. 2149 (Gasthörer Bd. 3: 1890-1904).
- [30] Schweizerisches Sozialarchiv, Bestand F.5153 Sonos – Schweizerischer Verband für Gehörlosen- und Hörgeschädigten-Organisationen, URL (30. 4. 2025): www.bild-video-ton.ch/bestand/signatur/F_5153/1000
- [31] Bibliothek Hör- und Sprachgeschädigtenwesen (Library for the Hearing and Speech Impaired) Leipzig, Sign. 1975/78.
- [32] BRAUN-ELWERT, RUDOLPH: personal communication, September 5, 2022.

THE GRENOBLE INSTITUTE OF PHONETICS AND ITS ROLE IN THE CONCEPTUAL TRANSFORMATION OF PHONETICS IN THE EARLY 20TH CENTURY

Coriandre Vilain¹, Nathalie Vallée¹

¹Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, 38000 Grenoble, France

* Institute of Engineering Univ. Grenoble Alpes

coriandre.vilain@gipsa-lab.fr, nathalie.vallee@gipsa-lab.fr

Abstract: The 19th century marked a scientific evolution in linguistics with the rise of experimental phonetics. Foreign language teaching gradually integrated this new discipline, leading to the creation of the first phonetics research laboratories. At the beginning of the 20th century, Grenoble's economy was booming and the university was expanding. Thanks to the *Comité de patronage des étudiants étrangers*, phonetics began to be taught. In 1904, a chair of modern French philology was created, entrusted to Théodore Rosset, who played a central role in establishing experimental phonetics in Grenoble. Rosset founded a phonetics laboratory on his arrival, gradually transforming it into an institute combining teaching and experimental research. He developed innovative teaching methods using phonographic recordings, speech vibration inscription and voice analysis instruments.

1 Introduction: the renewal of Phonetics in Europe at the end of the 19th century

The field of linguistic studies underwent a remarkable development in the 19th century, marked by the intellectual and cultural revolution of the Enlightenment and the legacy of the work of the founders of comparative grammar, such as E. Rask, W. von Humboldt, F. Bopp, A. Schleicher and G. von der Gabelentz. Inspired by the Lamarckian hypothesis, philologists adopted a natural conception of languages as evolving organic bodies, capable of adaptation and mutation [1]. In a paradigm encompassing homology and universalism, the historical study of languages turned towards heredity and the search for the natural causes of evolution at different depths: syntax, morphology of words, and naturalness of the sounds that express them [2].

1.1 From grammar to phonetic laws

From a comparative perspective, the description of sound changes in different eras led philologists to propose regularities in mutations. Towards the end of the 19th century, neo-grammarians formulated these regularities in terms of “laws” (like laws of physics), since they apply uniformly (even if there are exceptions, as long as these remain explicable) to the succession of stages in historical evolution. “Laws” mean that language evolution is not arbitrary, and that phonetic changes are predictable (e.g. [3]). The first phonetic laws issued by [4], and then [5], revealed the importance of the sound component and the rational interpretation of changes for understanding languages and their evolution, and establishing their kinship. There followed a preoccupation that was to accompany comparatists throughout the 19th century: the unsuitability of the Latin alphabet for transcribing the sounds of languages [6].

1.2 From phonetic laws to experimentation

To retrace the history of linguistic change, reasoning based on written forms alone proves insufficient: understanding sounds and the logic of their modification becomes necessary. For

this, a thorough knowledge of the sound elements of language is required. The ideas and principles of A. Comte's philosophical positivism, which spread in the second half of the 19th century, contributed to the development of this knowledge through new methodological approaches favoring observation, measurement and experimentation. These were made possible by the invention of new techniques and the creation of new instruments, such as those used by the physician and physiologist E. J. Marey to record and graphically inscribe speech [7]. Neogrammarians then sought the causes of phonetic changes in the physical, articulatory (mechanical actions) and acoustic (sonority analogies) characteristics of sounds [8]. Also contributing to the methodological renewal of the sciences was the dissemination of C. Bernard's experimental method, based on several key steps that alone enable us to move from the observation of a phenomenon to a causal and scientific explanation of phenomena [9].

1.3 The difficult path toward the recognition of oral language

Against this backdrop of renewed scientific knowledge, and soon after its creation in 1866, the *Société de Linguistique de Paris* (SLP) promoted the study of spoken language—oral rather than written—and, in 1874, set up an instrument commission at the initiative of its secretary, the philologist L. Havet [7]. However, at the end of the 19th century, the scientific approach to oral languages did not include regional languages and dialects (or patois), whose scientific interest in linguistics had not yet been clearly identified. The same applied to everyday language practices. As a result, although inspired by the scientific revival of the time, university institutions continued to give priority to the study of classical languages (Romance, Germanic, Slavic, etc.) from a historical and philological perspective, and to literary studies. Little attention was paid to the scientific study of dialects, which will continue to lack academic recognition for many years to come. This was despite early initiatives based on field surveys (e.g. those of J. Gilliéron, initiator with P. J. Rousselot in 1887 of the *Revue des patois gallo-romans*), of Rousselot himself [10], or a little later those of F. Brunot, instigator of the first phonographic surveys (1911-1914) and creator in 1911 of the *Archives de la parole* [11], [12], [13].

1.4 The development of practical foreign language instruction

In the first half of the 19th century, in connection with industrialization and the post-French Revolution era of 1789, education became institutionalized at the state level throughout Europe. To meet the needs of international exchange, the practical teaching of foreign languages was introduced in France as early as 1829 through the reform of Modern Education (ordinance by Henri de Vatimesnil). However, in its struggle with classical education, it was only under the ministry of Victor Duruy (1863–1869), and more decisively under the Third Republic, that it truly developed, driven by strategic national priorities [12], [14], [15]. Although the primacy of spoken language was included in ministerial guidelines as early as 1840 [14], phonetics remained virtually unknown until 1886 [16]: oral learning meant mastering a normative pronunciation model based on pronunciation manuals, always tied to written language [13], [17]. With P. Passy in France—then a young agrégation holder in English and founder in 1886 of the *Dhi Fonètik Tîtcierz' Asóciécon* (which became the *Association phonétique des professeurs de langue vivante* in 1889)—alongside others such as O. Jespersen in Copenhagen and H. Sweet in London, phonetics entered universities as the science of spoken sounds.

Scientific advances allowed for the development of a modern approach to language teaching, integrating into “natural” learning methods new knowledge from the physics and physiology of speech sounds [12], [18]. The success of these new pedagogical methods—particularly in the context of growing international exchanges—soon led to the creation of university training programs in the form of workshops or summer courses for foreign students wishing to improve their French for professional or cultural purposes. The University of Grenoble was among the first to launch such programs, as early as 1896 [12], [19]. Between the late 19th and early 20th centuries, the first experimental phonetics laboratories began to spread across France [20],

[21]: Paris in 1897 (P. J. Rousselot), Grenoble in 1904 (T. Rosset), Montpellier in 1905 (M. Grammont), Lille in 1911 (F. Piquet)... and internationally: Kazan in Russia in 1895 (V. A. Bogorodickij), Hamburg in 1910 (G. Panconcelli-Calzia), London in 1912 (D. Jones), Barcelona in 1913 (P. Barnils), Perth, Australia in 1913 (E. Suddard). Despite their geographic spread, phonetics laboratories remained relatively few in number, and some evolved into full-fledged institutes. Why did the University of Grenoble become home to one of the first phonetics institutes in France and Europe [22: 24], and how did this institute position itself within the legacy of P. J. Rousselot and E. W. Scripture?

2 The emergence of Phonetics in Grenoble

2.1 Grenoble at the turn of the 20th century

At the beginning of the 20th century, Grenoble was a dynamic city whose population had been steadily increasing since the early 19th century [23]. Between 1806 and 1901, it grew from 22,000 to 69,000 inhabitants. This demographic growth can be explained by several factors. First, the economic dynamism of the city and its region, which saw the rise of hydroelectric industries—referred to as “*houille blanche*” (*white coal*)—such as paper mills, boiler-making, electrometallurgy, electrochemistry, and others. These industries attracted thousands of workers to Grenoble, both from within France and abroad (notably from neighboring Italy). In addition to this economic vitality came the development of mountain tourism. The Isère branch of the French Alpine Club was founded in 1874 by H. Duhamel and C. Lory, a professor of geology at the University of Grenoble. In 1889, the *Syndicat d’initiative de Grenoble et du Dauphiné*—the first tourist office of its kind in France—was established.

This economic and touristic dynamism in the early 20th century led to the development of the University of Grenoble. Although it is one of the oldest universities in France (founded in 1339 by Humbert II, the last ruler of the Dauphiné), the University of Grenoble had experienced a chaotic history over the centuries, constantly competing with the neighboring universities of Vienne and Valence [24]. However, it found a passionate advocate in J. Fourier—an eminent mathematician and prefect of the Isère department from 1802 to 1815—who actively supported the *Université impériale*, created in 1806. Fourier developed the three faculties of Law, Humanities, and Science. While Law had been a central discipline in Grenoble since the university’s foundation (Dauphin Humbert II had a keen interest in Roman law), the Humanities and Sciences truly began to flourish under the *Université impériale*. Figures such as J. Fourier, the Champollion brothers (including Jean-François, the decipherer of hieroglyphs), and the botanist Dominique Villard played key roles in the intellectual life of Grenoble at the time.

However, in the early 19th century—just as in the rest of the academic world—there was no mention of teaching or research in phonetics. It was only toward the end of the 19th century that interest in this new discipline, experimental phonetics, began to grow within the Faculty of Humanities. This development was initiated by the creation in 1896 of the “*Comité de patronage des étudiants étrangers*” (Committee for the Support of Foreign Students) at the University of Grenoble, led by M. Reymond [25].

2.2 The *Comité de patronage des étudiants étrangers* (CPEE)

The CPEE of Grenoble was a group of academics, industrialists, and prominent local figures who came together in 1896 under the initiative of M. Reymond to promote Grenoble and its region among foreign students. The creation of a CPEE in Grenoble was not an isolated case in France; several such committees existed at the end of the 19th century. All appear to follow the model of the Paris committee, founded in 1891 and directed by P. Melon.

The statutes of the Grenoble CPEE specify in their first article¹:

*“La société qui prend le nom de comité de patronage des étudiants étrangers près l’université de Grenoble, fondée le 3 juillet 1896, a pour but d’inviter par la plus large publicité possible les étrangers à venir faire leurs études à Grenoble, de faciliter leur installation ; de leur donner un appui moral et d’organiser les enseignements les plus propres à attirer une clientèle étrangère, soit pendant les vacances, soit pendant le cours de l’année scolaire.”*²

The early days of the CPEE in Grenoble were modest. In 1896, only one foreign student—a German—enrolled at the university. But very quickly, the number of foreign students of various nationalities began to grow [19]. The CPEE notably funded university teaching positions and hosted visiting professors to give seminars during the holiday sessions. Thus, several eminent linguists came to teach phonetics during the summer courses: M. Bréal in 1899, P. J. Rousselot in 1900, and P. Passy in 1903. In 1904, in response to the growing success and increasing number of foreign students coming to Grenoble to learn French, the CPEE began discussions with the Faculty of Letters to propose the creation of a lectureship in modern French philology for the teaching of French to foreign students. The job announcement was published in 1904, and the Faculty Assembly reviewed applications on June 27, 1904. Ultimately, T. Rosset was selected from among five candidates on November 1, 1904. His responsibilities included 1 hour of modern French phonetics, 2 hours of grammar (morphology and syntax), and 1 hour of text commentary. Upon his arrival as Lecturer, Rosset set about developing his laboratory, attached to the Faculty of Letters. As he recalled in 1909, the laboratory’s initial funding came directly from the Ministry. Since the University of Grenoble was unable to provide a location for the laboratory, the CPEE turned to the city of Grenoble. In May 1905, an abandoned room in a primary school was equipped with two tables, a phonograph, and a portable recorder. This marked the beginning of the phonetics laboratory in Grenoble [26].

2.3 Théodore Rosset

At the time volumes 1 and 2 of *Principes de phonétique expérimentale* by Rousselot were published (1887, 1901), Rosset was serving in the military. He was 21 years old. His military record indicates that he was a student residing at *Lycée Michelet* in Paris and pursuing a degree in Literature, which exempted him from the final two years of national service in exchange for a ten-year reserve commitment and continued studies. From 1901 to 1904, Rosset was a full student in Section IV (Historical and Philological Sciences) at the *École des Hautes Études* (EHE), most likely admitted on the recommendation of a faculty member after a short probationary period to assess his interest in research. His professors (or research directors) in philological sciences included L. Havet in Latin Philology, M. Bréal and A. Meillet in Comparative Grammar, G. Paris, A. Thomas, and M. Roques in Romance Philology, J. Gilliéron in the Dialectology of Roman Gaul, and P. Passy in General and Comparative Phonetics. Although many philologists were interested in phonetics and supported Rousselot’s experimental method, few actually practiced it themselves [27], the experimental method was not part of the EHE’s curriculum. Even though some syllabi mentioned field visits to museums or courses held at the Louvre (e.g., Assyrian Philology, Roman Epigraphy, Literary History of the Renaissance), there is no indication that Rousselot’s laboratory was included in the phonetics program. The phonetics

¹ Statuts du Comité de Patronage des Étudiants Étrangers près l’université de Grenoble, 1902.

² The society, which adopts the name Committee for the Patronage of Foreign Students at the University of Grenoble, founded on July 3, 1896, aims to encourage foreigners to come study in Grenoble through the widest possible publicity, to facilitate their settlement, to provide them with moral support, and to organize the types of courses most likely to attract a foreign clientele, either during holidays or throughout the academic year. [our translation]

taught at the EHE followed the “classical” scientific methods of the Neogrammarians. According to [11] and [17], this absence can be explained by: i) Linguists’ scepticism toward the experimental approach, which required new technical skills that many were reluctant to acquire—as Passy himself would admit in 1932 [17]; ii) The fear of marginalization or exclusion from the academic world for deviating from the scientific methods of Neogrammarian philology [11].

In the same year he earned his licence ès Lettres, Rosset passed the *agrégation* in grammar in 1904, ranking 4th. The director of *agrégation* training at the Sorbonne at that time was F. Brunot, under whose supervision Rosset conducted the research that would result in his first publications [28], [29]. But his real entry into the field of experimental phonetics for French language teaching came with the publication of *Exercices pratiques d'articulation et de diction composés pour l'enseignement de la langue française aux étrangers* [30]. He threw himself energetically into developing this course while becoming increasingly involved in the CPEE’s activities, eventually becoming its secretary in 1907.

2.4 The birth of the Phonetics Institute of Grenoble

Benefiting from an increasingly substantial budget—mainly from the registration fees paid by foreign students and royalties from the sales of Rosset’s *Exercices pratiques*—the phonetics laboratory gradually expanded. An inventory written by Rosset in 1911 of the equipment, furniture, and books he had acquired since the laboratory’s inception allows us to trace its development. By 1911, we find a heterogeneous list including instruments such as a Koenig analyzer, a Lumière cinematograph, Edison phonographs, etc.; as well as laboratory supplies (oil, kerosene, gasoline, alcohol), test tubes, rubber tubing. With its growing equipment and increasing activity, the laboratory moved in May 1908 to a more “respectable” room than the outdated space of its early days (figure 1). At this point, it gained the status of *Institute of Phonetics*, combining research and university teaching. This new status gave the phonetics program in Grenoble a unique place in the French and European academic landscapes [22].

“Il y a dans la plupart des Facultés des Lettres un cours de phonétique historique ; il y a à l'École des Hautes Études, à Paris, et dans diverses Universités étrangères des séminaires de phonétique descriptive où l'on décrit et enseigne aux futurs professeurs de langues vivantes les sons de ces langues ; il y a enfin, au Collège de France et dans cinq ou six Universités françaises et étrangères, des laboratoires de phonétique expérimentale ; mais nulle part on n'a songé à réunir en un seul corps ces tronçons épars des études phonétiques. C'est une création originale de l'Université de Grenoble que d'avoir groupé en un Institut particulier toutes ces recherches pratiques, historiques, dialectales et expérimentales, et attesté par là leur intime solidarité.”³ [30: 445]

As for the research activities mentioned by Rosset, while the practical, historical, and experimental aspects are clearly attested, it is more difficult to find evidence of dialectal research in his publications. Yet, Rosset noted—in 1909—that “*L'Institut est membre fondateur de la Société internationale de dialectologie romane et il sera officiellement chargé de faire l'étude*

³ Most Faculties of Letters have a course in historical phonetics; there are descriptive phonetics seminars at the *École des Hautes Études* in Paris and in various foreign universities, where the sounds of Modern languages are taught to future language teachers; there are, finally, experimental phonetics laboratories at the *Collège de France* and in five or six French and foreign universities. But nowhere has anyone thought to bring together these separate branches of phonetic studies into a single body. It is an original creation of the University of Grenoble to have grouped all these practical, historical, dialectal, and experimental studies into a distinct Institute, thereby attesting to their intimate interconnection. [our translation]

des patois et dialectes du Dauphiné”⁴ [30: 449]. In contrast, Rosset’s activity as an engineer and educator is strikingly illustrated by his complementary thesis [31].

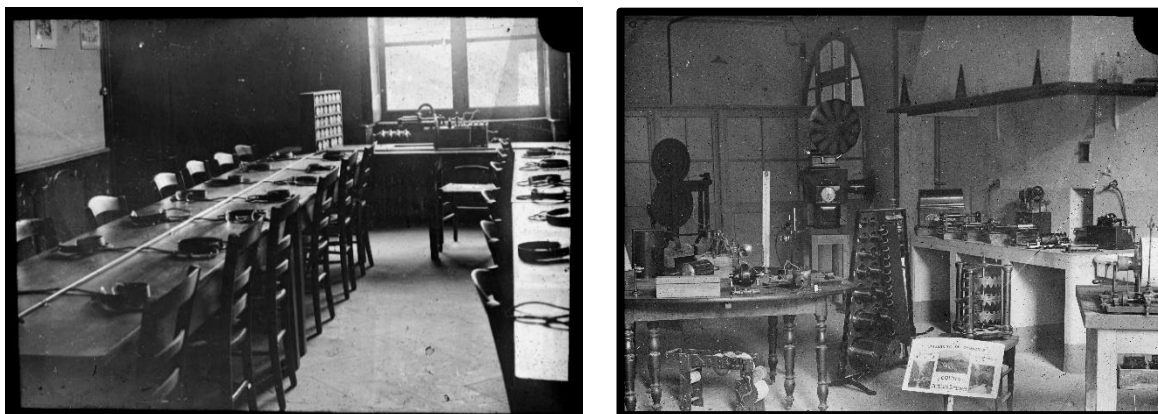


Figure 1. The language laboratory (left) and the phonetics laboratory (right) of the Grenoble Institute of Phonetics around 1908

3 The Theodore Rosset’s methodology and its dissemination

As Rosset reports in the preface of [30], there were already many French language teaching manuals that presented the theoretical foundations of French pronunciation (he cites Passy, 1899; Rousselot and Laclotte, 1902; Zünd-Burguet, 1902; Nyrop, 1902). These works, though at the forefront of phonetic knowledge, according to him “merely apply it through the reading of random texts simply strung together” [30: 2]. In his view, however, a new method was needed—one that extended the work of Rousselot and Zünd-Burguet and applied it directly to the learning of French:

*“C’est de leurs conclusions que s’inspire cette méthode. À côté de l’enseignement théorique, on veut mettre désormais la démonstration expérimentale des articulations; le palais artificiel, les ampoules exploratrices, le cadran indicateur, le cylindre inscripteur, le tambour enregistreur, le manomètre à eau, le signal du larynx, etc., permettent désormais de connaître et de montrer exactement quels organes interviennent dans la production du son, dans quelle mesure, à quel moment ; ils peuvent aussi révéler quelles erreurs commet un étranger dans la mise en action des organes phonateurs ; ils lui permettent de se rendre compte lui-même, par la vue, que a allemand ne s’articule pas comme a français, de vérifier expérimentalement si les corrections qu’il essaye sont heureuses, de s’assurer enfin qu’il met bien en mouvement les organes nécessaires, ceux-là seulement et dans la mesure exacte qui convient.”*⁵ [30: 2–3]

In this way, experimental phonetics could—in Rosset’s view—serve learners by allowing

⁴ The Institute is a founding member of the International Society for Romance Dialectology and will be officially tasked with studying the patois and dialects of the Dauphiné. [our translation]

⁵ This method is inspired by their conclusions. Alongside theoretical instruction, we now aim to provide the experimental demonstration of articulations; the artificial palate, exploring bulbs, indicator dial, recording cylinder, kymograph drum, water manometer, larynx signal, etc., now make it possible to determine and demonstrate precisely which organs are involved in producing a sound, to what extent, and at what moment; they can also reveal what errors a foreigner makes in activating the phonatory organs; they allow the learner to see for themselves, visually, that the German *a* is not articulated like the French *a*, to experimentally verify whether their corrections are successful, and to ensure that they are using the appropriate organs, and only those, with the correct degree of activation. [our translation]

them to measure and visualize their own speech production. But learning a language does not only consist in articulating isolated words; one must also grasp its prosody:

“De plus, parler une langue correctement, ce n'est pas articuler sans fautes des mots isolés, c'est prononcer des phrases avec l'accent, les accommodations, le rythme, l'intonation qu'un indigène leur donne spontanément, et qu'un étranger doit apprendre, avec peine parfois. Le plus souvent on confie cet enseignement à la vie quotidienne ; et l'on pense qu'à force d'entendre parler et lire, on arrivera à parler et à lire correctement. Cet espoir est parfois réalisé mais une telle méthode demande beaucoup de temps, et le succès en est toujours incertain”⁶ [30: 2–3]

To address this, nothing beats a dedicated teacher, “constantly available to students to tirelessly repeat the same ten lines of prose” in a consistent manner. To that end, Rosset proposed the mechanical reproduction of sounds recorded on phonograph cylinders. His *Exercices pratiques* were thus accompanied by recordings of Rosset himself pronouncing the texts included in the manual. The phonograph in the phonetics laboratory thus served as a “tireless teacher,” and Rosset set out to develop a device capable of duplicating phonograph cylinders—which quickly wore out with use—and added to it a mechanism for visualizing recorded speech signals. This was his principal contribution to experimental phonetics:

“Aussi, dès le début, ai-je été préoccupé de posséder un moyen d'obtenir mécaniquement la copie d'un cylindre enregistré sur un cylindre vierge, qui remédierait à tous ces inconvénients et qui, de plus, permettrait à mes élèves étrangers d'emporter chez eux en même temps un livre pour avoir la transcription phonologique des sons français et une collection de cylindres pour avoir la transcription phonographique de la diction française”⁷ [31: 63]

The technical challenges were many, and it was with the help of a “mechanic” that he finally managed to build the duplicating device, after numerous trials detailed in his manuscript. The task became even more difficult when Rosset sought to add to his setup a means of recording the signal onto paper for precise measurement and analysis. Taking inspiration from L. Hermann, he proposed an optical system based on a small mirror connected to the phonograph's playback head, which reflected a beam of light onto a cylinder wrapped in photographic paper and housed in a darkroom. This system allowed him to amplify the recorded signal without the distortions caused by a mechanical lever [32] (figure 2). The signals produced by this device—presented in the appendix to Rosset's thesis—were remarkably clear, but their relevance was quickly questioned by leading phoneticists of the time, including Rousselot or Poirot [33], [34]. However, Chlumský's opinion about the reliability of Rosset's new technical proposal was less critical [35], [36]. Rosset also had Brunot's unwavering support:

“Mais F. Brunot était fermement convaincu de cette supériorité [de la méthode Rosset]. Lors de l'inauguration des Archives de la parole (Paris, 3

⁶ Moreover, speaking a language correctly is not merely about pronouncing isolated words without error—it is about speaking in phrases with the stress, adjustments, rhythm, and intonation that a native speaker naturally employs, and that a foreigner must learn, sometimes with difficulty. Most often, this instruction is left to daily life; it is assumed that by hearing enough spoken and read language, one will eventually speak and read correctly. This hope is sometimes fulfilled, but such a method takes a great deal of time, and its success is always uncertain [our translation]

⁷ From the very beginning, I was intent on obtaining a way to mechanically copy a recorded cylinder onto a blank one, which would remedy all these inconveniences and, moreover, allow my foreign students to take home with them both a book with the phonological transcription of French sounds and a collection of cylinders with the phonographic transcription of French speech [our translation]

juin 1911), Brunot indiqua comme l'un des buts poursuivis par cette initiative « la création définitive d'une nouvelle science : la phonétique » et ajoutait « Je suis de ceux qui espèrent depuis vingt ans qu'elle va enfin se constituer ». A cette occasion, il illustra par des projections lumineuses la supériorité incontestable de l'appareil et des tracés de Rosset qui enfermaient en eux, au dire de Brunot, l'avenir de la phonétique.”⁸ [18: 51]

Despite the criticism expressed by some of the recognized experts of the time, the concept of a language laboratory according to the Rosset method combining theoretical research in phonetics with advanced experimental technologies for the analysis and remediation of foreign language production and audio-visual perception, has spread beyond the borders of France: in England [37], Italia [38], United States [19] and Abou Haidar (this volume)... There is no doubt that he trained or influenced several generations of teachers of French as a foreign language, including abroad, through his lectures, publications and training. Its publications, mainly in French, may have been a hindrance to the dissemination of his method outside the circle of phoneticians who teach French as a foreign language.

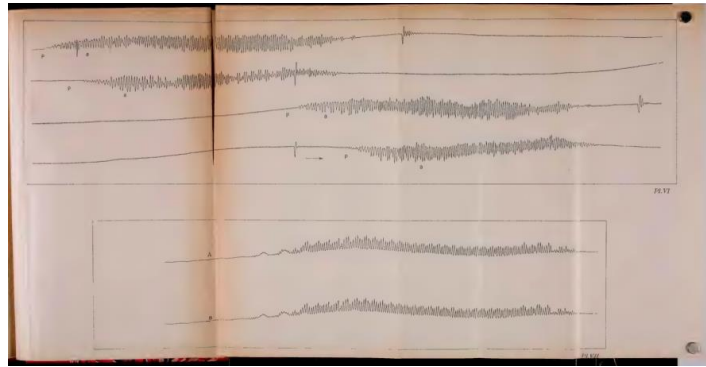


Figure 2. T. Rosset and its apparatus around 1910 (left), and an example of amplified recorded signal (right)

4 Conclusion

Undoubtedly, Rosset contributed to including language learning in the new phonetic science in the sense of A. Comte, C. Bernard and E. W. Scripture in which experiments constitute the only possible approach to objectify explanations of what is observed.

*“La phonétique expérimentale permettra dans l'avenir d'avoir des idées précises sur les transformations des articulations, sur les altérations des sons qui en résultent et probablement de préciser les causes et les moyens de ces transformations”*⁹ [31: 18 footnote 1]

⁸ Nonetheless, F. Brunot was firmly convinced of the superiority of Rosset's method. At the inauguration of the *Archives de la Parole* in Paris on June 3, 1911, Brunot identified among the aims of the initiative “the definitive establishment of a new science: phonetics,” further declaring, “I am one of those who have hoped for twenty years that it would finally come into being.” On this occasion, he employed lantern projections to demonstrate the undeniable superiority of Rosset's device and graphical tracings, which, in Brunot's view, embodied the very future of phonetic science. [our translation]

⁹ Experimental phonetics will in the future allow us to have precise ideas on the transformations of articulations, on the resulting alterations of sounds and probably to specify the causes and means of these transformations [our translation]

In particular, from the very beginning of the Grenoble phonetics laboratory, Rosset was keen to demonstrate the importance of inscription methods of speech curves in order to be able to carry out detailed analyses of the productions of foreign language learners [39].

From one perspective, language learning gained scientifically grounded methodologies through Rosset and the development of his Institute of Phonetics. From another, phonetic theory itself benefited from the quasi-experimental implementation of phonetic instruction, anticipating—only a few years later—the emergence of key theoretical constructs such as phonological representation and the complex cognitive mechanisms involved in the processing of non-native speech sounds. This includes, in particular, the notion of phonological deafness in foreign language acquisition [40: 79–80] and the concept of the phonological sieve [41: 54].

References

- [1] BOPP, F. 1827. *Ausführliches Lehrgebäude der Sanskrita-Sprache*. Berlin: Druckeri der Königl. Akademie der Wissenschaften.
- [2] SAMPSON, G. 1980. *Schools of linguistics*. Stanford, CA: Stanford University Press.
- [3] PASSY, P. 1891. *Étude sur les changements phonétiques et leurs caractères généraux*. Paris: Firmin-Didot.
- [4] RASK, R. 2013 [1818]. *Investigation of the origin of the Old Norse or Icelandic language*. New edition of the 1993 English translation by Niels Ege, with an introduction by Frans Gregersen. Amsterdam & Philadelphia: John Benjamins.
- [5] GRIMM, J. 1822. *Deutsche Grammatik*. Göttingen: Die Dieterische Buchhandlung.
- [6] ROUSSEAU, J. 1981. R. Rask (1787-1832) et la transcription des langues amérindiennes – Une lettre inédite à J. Pickering. *Histoire Épistémologie Langage* 3-2. 69–83.
- [7] TESTON, B. 2010. Le rôle d'Étienne-Jules Marey dans l'émergence de la phonétique expérimentale. In L.-J. Boë & C. E. Vilain (eds), *Un siècle de phonétique expérimentale*, 59–89. Lyon: ENS Éditions.
- [8] PAVEAU, M.-A. & G.-E. SARFATI. 2003. *Les grandes théories de la linguistique ; de la grammaire comparée à la pragmatique*. Paris: Armand Colin.
- [9] BANGE, C. 2009. Claude Bernard, la méthode expérimentale, et la Société de Biologie. *Journal de la Société de Biologie* 203(3). 235–247.
- [10] ROUSSELOT, P. J. 1891. *Les modifications phonétiques du langage étudiées dans le patois d'une famille de Cellefrouin (Charente)*. Paris: Welter.
- [11] BERGOUNIOUX, G. 1992. Linguistique et variation : repères historiques. *Langages* 108. 114–125.
- [12] CHEVALIER, J.-C. 1997. Phonétique expérimentale et Cours pour étrangers en France à la fin du 19^e. *Documents pour l'histoire du français langue étrangère ou seconde* 19. 158–165.
- [13] WEBER, C. 2019. Interrogations épistémologiques autour de l'oralité. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle* 16(1).
- [14] PUREN, C. 1989. L'enseignement scolaire des langues vivantes étrangères en France au XIX^e siècle ou la naissance d'une didactique. *Langue française* 82. 8–19.
- [15] DUBOIS, J. 2017. L'enseignement des langues étrangères sous la Troisième République : des disciplines en prise avec les relations internationales. *Revue française de pédagogie* 199. 23–37.
- [16] PASSY, P. 1910. Un quart de siècle. *Le Maître Phonétique*. 69–72.
- [17] GALAZZI, E. 1995. Phonétique/Université/Enseignement à la fin du XIX^e siècle. *Histoire épistémologie langage* 17(1). 95–114.
- [18] GALAZZI, E. 1997. Physiologie de la parole et enseignement de la prononciation fin 19^e/début 20^e siècles. *Documents pour l'histoire du français langue étrangère ou seconde* 19. 166–183.
- [19] ABOU HAIDAR, L. 2024. Ce que la didactique doit à la phonétique : l'hypothèse du chaînon

- manquant. *Repères-Dorif : autour du français : langues, cultures et plurilinguisme* 30.
- [20] BOË, L.-J., & J.-F. BONNOT 2010. Création et émergence de la phonétique expérimentale en France : de l'abbé Rousselot à Théodore Rosset. In L.-J. Boë et C. E. Vilain (eds.), *Un siècle de phonétique expérimentale*, 103–128. Lyon: ENS éditions.
- [21] BROCK, G., C. BUFFONI, J. CLARENC, P.-O. GAUMIN & F. HIRSCH. 2021. Les laboratoires de phonétique français dans la première moitié du XX^e siècle : instrumentations, corpus et thématiques de recherche. In C. Dodane & C. Schweister (eds), *Histoire de la description de la parole : de l'introspection à l'instrumentation*, 301–318. Paris: Honoré Champion.
- [22] CUQ, J.-P. 1997. La chaire de philologie moderne destinée à l'enseignement du français aux étrangers à l'Université de Grenoble. Création et évolution : 1904-1936. *Documents pour l'histoire du français langue étrangère ou seconde* 20. 22–36.
- [23] SYLVESTRE, V. 1926. La population de Grenoble en 1926. *Revue de géographie alpine* 14(3). 659-664.
- [24] FAVIER, R. 2017. *Le roman de l'université Grenoble, 1339-2016*. Grenoble: Presses universitaires de Grenoble (PUG).
- [25] RAKOTOFIRINGA, H. 1997. Contributions du Comité de Patronage des Étudiants Étrangers au développement de l'Université de Grenoble durant les premières décennies de ce siècle. *Documents pour l'histoire du français langue étrangère ou seconde* 20. 73–85.
- [26] ROSSET, T. 1909. Les agrandissements de l'université de Grenoble – 1/ L'institut de phonétique. *Annales de l'Université de Grenoble* 21. 444–449.
- [27] GALAZZI, E. 2014. Pierre Jean Rousselot : la phonétique expérimentale au service de l'homme. *Dossiers d'HEL, Linguistiques d'intervention : des usages socio-politiques des savoirs sur le langage et les langues* 6.
- [28] ROSSET, T. 1904. E féminin au XVII^e siècle. In *Mélanges de philologie offerts à Ferdinand Brunot à l'occasion de sa 20^e année de professorat dans l'enseignement supérieur par ses élèves français et étrangers*, 433–450. Paris: Société nouvelle de librairie et d'édition.
- [29] ROSSET, T. 1908. *Entretien, Doutes, Critique et Remarques du Père Bouhours sur la Langue Française, 1671-1692*. Grenoble: Allier Frère.
- [30] ROSSET, T. 1905. *Exercices pratiques d'articulation et de diction*. Grenoble: Gratier.
- [31] ROSSET, T. 1911. *Recherches expérimentales pour l'inscription de la voix parlée*. Paris: Armand Colin.
- [32] TESTON, B. 2006. À la poursuite de la trace du signal de parole. In *Proc. XXVI^{es} Journées d'Études sur la Parole, JEP-AFCP*. Dinard, France. 7–10.
- [33] ROUSSELOT, P. J. 1911. Critique d'expériences. *Revue de phonétique* 1. 201–231.
- [34] POIROT, J. 1911. Compte-rendu des Recherches expérimentales pour l'inscription de la voix parlée. *Revue de phonétique* 1. 190–196.
- [35] CHLUMSKÝ, J. 1911. Appareils nouveaux. *Revue de phonétique* 1. 68–78.
- [36] CHLUMSKÝ, J. 1911. Compte rendu. *Revue de phonétique* 1. 376–383.
- [37] JONES, D. 1909. fonetics et *Grønøbl. *Le Maître Phonétique* 11-12, 143–146.
- [38] GALAZZI, E. 2002. Echos phonétiques en Sicile entre XIX^e et XX^e siècles. *Documents pour l'histoire du français langue étrangère ou seconde*, (28), 13–31.
- [39] ROSSET, T. (1905). Un enseignement expérimental de la prononciation française. *Revue internationale de l'enseignement* 49. 310-315.
- [40] POLIVANOV, E. 1931. La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4. 79–96.
- [41] TROUBETZKOY, N. S. 1967 [1939]. *Principes de phonologie*. (2nd ed). Paris: Klincksieck.

SOME HERITAGE DOCUMENTS FROM HERITAGE INSTRUMENTS AT THE PHONETICS INSTITUTE OF STRASBOURG

Béatrice Vaxelaire¹, Rémi Yirméyan Belem¹, Anastasia Kananovich¹, Jinwoo Cha¹, Mariama Jobe¹, Lisa Matias-Gaspar¹, Emma Tasset¹, Noé Xiu¹, Fabrice Marsac¹, Rudolph Sock^{1,2}

¹University of Strasbourg, Institut de Phonétique, de Philologie et de Linguistique de Strasbourg, UR Linguistique, Langues et Parole – LiLPa, France,

^{1,2}Pavol Jozef Šafárik University, Language Information and Communication Laboratory – LICOLAB, Slovakia

sock@unistra.fr

Abstract: The main tack of this paper is to pinpoint the usefulness of some of the many old instruments and documents that have been serving efficaciously at The Phonetics Institute of Strasbourg as teaching tools in experimental phonetics, since the 40s. The authors have deliberately not tried to improve the quality of these old documents, by adapting them, in order to sensitise the reader to their relevance, then and today, for understanding phonetic phenomena, however complex they may be. Indeed, when heritage scientific instruments and documents, like those that come from the past in the area of experimental phonetics, are used and interpreted appropriately, they help to underscore the findings of early phoneticians, which comprised the principles of reflection, design and creation of adapted instruments, hypothesis formulation, and empirical data, today keystones of contemporary phonetic method.

1 Introduction

Georges Straka founded The Phonetics Institute of Strasbourg in 1947. The originality of the institute lay in its commitment to the experimental method and the development of devices for acquiring articulatory and acoustic data, in particular radio-cinematography, designed to describe articulatory physiology in speech production.

In addition to radio-cinematography that is the hallmark of the institute, the laboratory has successively used mechanical (kymograph, palatograph, etc.) and electronic (spectrograph, sonagraph, mingograph, etc.) devices, totalling around one hundred in all. Construction engineers of laboratory equipment mainly assumed building these devices. However, some instruments were made-in-home devices, the work being carried out, as from the early 60s, by laboratory technicians, like Gilbert Brock, Paul Burgstahler, Michel Massuelle and André Zorn, to name some of them, supervised by Georges Straka and Péla Simon.

The institute still uses heritage instruments and documents for didactic purposes. Even if the quality of these documents is not always optimal, as we can see by going through this article, they still serve to illustrate the teaching content the phonetics teachers at the institute desire to convey to the students and, on some occasions, to the public at vulgarisation events.

As a final remark, we should underscore that the heritage documents produced by these heritage articulatory and acoustic data acquisition instruments have been essential to current studies in speech production and perception. Indeed, modern high performance acquisition systems, which provide high quality digital documents, have relied heavily on past

knowledge. The expertise is not as new as it seems; it is like putting “old wine in a new bottle”.

2 Palatograms

Palatograms are graphic representations of the area of the palate in contact with the tongue during production of specific speech sounds. Such documents were acquired, as from the early 60's at The Phonetics Institute of Strasbourg by a mechanical device, the static Palatography (see Figure 1, a made-in-house palatography), which was the method that consisted of brushing the subject's tongue with a mixture of natural charcoal and dark chocolate powder. The tongue print left on the palate during articulation was then captured using a mirror and a camera (see Moses [1], for a brief history of palatography).

Figures 1a, 1b and 1c are palatograms that show prints of the manner and place of articulation of two lingual plosives and one lingual constrictive: the alveolar-dental /t/, the dorsal-velar /k/ and the alveolar constrictive /l/ in the utterance /tap/ for the French words “tape” (tap), /kap/ “cape” (a cape) and /palpa/ “palpa” (palpated), respectively. Whereas all three sounds indicate contact on the medial sagittal axis of the hard palate, the /t/ contact is in the front region, the /k/ in the rear region and the /l/ print is limited to the alveolar region, compared with that left by the /t/ contact, which spreads to the upper incisors.



Figure 1. A made in-house palatography at The Phonetics Institute of Strasbourg around the 1960.

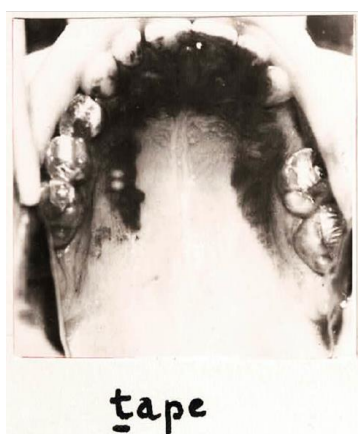


Figure 1a. Palatogram of /t/



Figure 1b. Palatogram of /k/

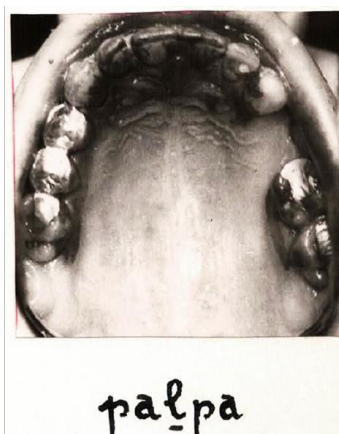


Figure 1c. Palatogram of /l/

The palatograms in Figures 1d and 1e show constrictions in the post-alveolar region for /ʒ/ and in the palatal region for /j/ of the vocal tract in the utterances /ʒɑ̃/ for the French word “Jean” (John) and /jɑ̃b/ “iambe” (iamb), respectively. The smaller width necessary for the production of the voiced apical post-alveolar fricative /ʒ/ is responsible for the characteristic

turbulent noise in the acoustic spectrum due to frication, compared with the much larger constriction in the dorsal-palatal /j/ constrictive, resulting in the absence of turbulence in the acoustic spectrum.

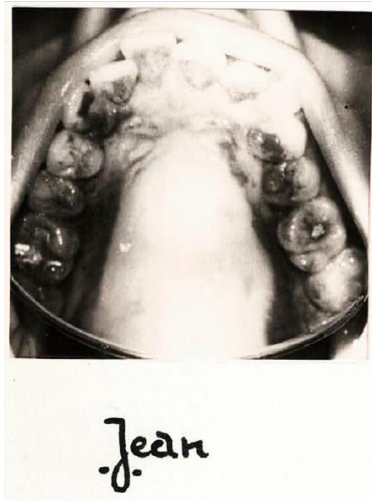


Figure 1d. Palatogram of /ʒ/

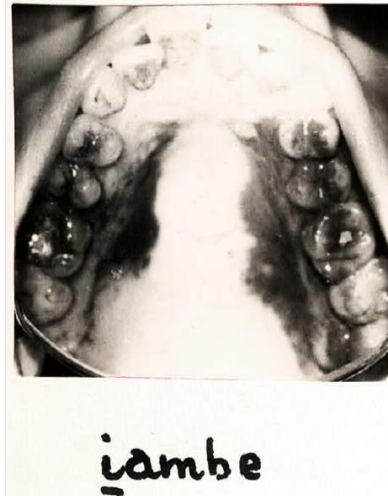


Figure 1e. Palatogram of /j/

3 Kymograms

Kymograms are graphic tracings, which illustrate phonetic (acoustic and aerodynamic) events, as a function of time. Kymograms were obtained at The Phonetics Institute of Strasbourg using a device called the Kymograph (see Chafcouloff [2] for a detailed description of this device). The kymograph (Figure 2), a machine based on a mechanical principle, was made up of two parts: 1) a “recording” part comprising a system for capturing the flow of phonatory air from the mouth and nose, via a mouthpiece for mouth pressure and one or two glass olives for nasal pressure. The airflow is channelled through rubber tubes to the membranes, which vibrate under the effect of the air pressure, thus obtaining resonances of airflow from the mouth.

At the same time, laryngeal vibrations were recorded by applying a diaphragm to the larynx at the level of the vocal folds; 2) an “inscription” part, the main element of which is a cylinder (Marey’s drum), itself driven by a mechanical rotation system, wrapped with smoked paper, upon which the waves were recorded. For a description of the original version of the kymograph, the reader is referred to abbé Rousselot [3], and for a much later version, the “polyphonometer”, to Teston [4].



Figure 2. A photo showing the phonetic kymograph as supplied to The Phonetics Institute of Strasbourg, together with its accessories by G. Boulitte around 1947.

Figures 2a and 2b are kymograms that show incomplete voicing of the voiceless plosives /k/, /p/ flanked by vowels in the utterance /səluqikipø/ (Figure 2a) in the French phrase “celui qui peut” (he who can) and of the voiceless fricative /s/ in the sequence /dãsõ/ (Figure 2b) within the phrase “dans son Paris” (in his/her Paris).

From top: “L” stands for laryngeal vibrations, “N” for nasal air pressure, “Ba” for mouth air pressure and “Bm” for mouth resonances, different from an acoustic pressure wave, such as that obtained from a microphone.

In the /iki/ sequence (Figure 2a), delayed voicing decay from vowel /i/ (“fin de la sonorité”), visible in the laryngeal trace (“L”), can be observed after onset of closure for the plosive /k/ (“début de k”), revealed by a significant drop in mouth air pressure (“Ba”). The scenario is similar in the /ipø/ context. Figure 2b shows voicing in the laryngeal trace (“L”) within the /s/ fricative interval (“fin de la sonorité”), after offset of mouth resonances (“Bm”), and a remarkable rise in mouth air pressure (“Ba”) necessary for the production of the fricative, annotated as “début S”. These are all examples of partial carryover assimilation.

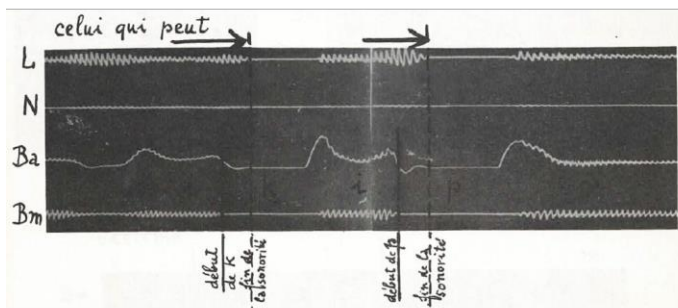


Figure 2a. Kymogram of /səluqikipø/

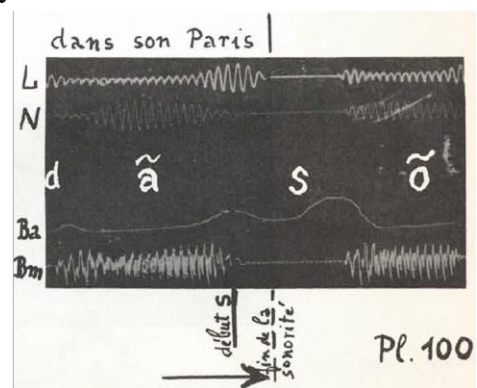


Figure 2b. Kymogram of /dãsõ/

Figure 2c is a typical kymogram showing incomplete devoicing of the vibrant dorsal-uvular constrictive /R/ in the utterance /partu/, corresponding to the French word “partout” (“everywhere”). Partial devoicing is provoked by the voiceless plosive /t/. This anticipatory assimilation is depicted in the laryngeal trace (“L”), and in the mouth resonances (“Bm”). Indeed, the very few pulses of the waves for the larynx and the mouth (after the left vertical reference line in the figure) within the /R/ interval, which consist mainly of the fundamental (first harmonic) of the voice (Ashby & Ashby [5]), are of very low amplitudes. The two or three flaps, corresponding to tongue dorsum and uvular contacts, are visible in the mouth air pressure trace (“Ba”).

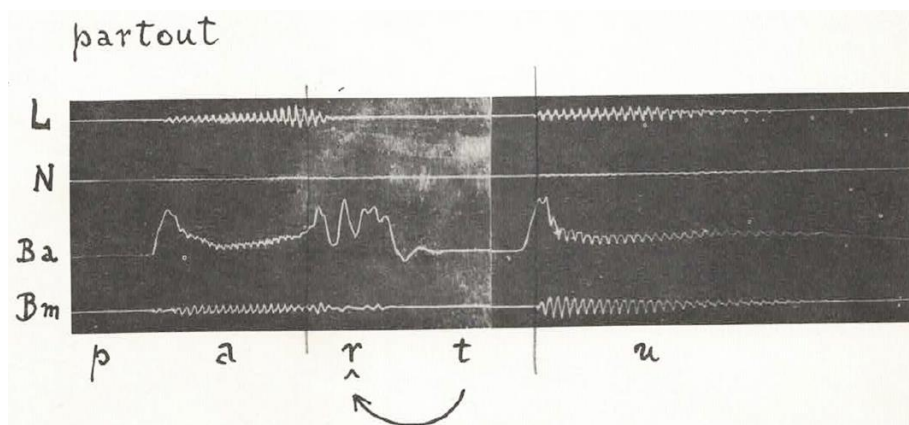


Figure 2c. Kymogram of /partu/

4 Mingograms

4-Line or 4-Channel Mingograms are plots that show, as the name suggests, four channels of graphical representations of acoustic and aerodynamic activity at the laryngeal, nasal and oral (resonances and pressure) levels. These documents were obtained using an electronically designed Mingograph, another made-in-house device, which was a derived and relatively transportable version of the Kymograph. The ‘Larynx-microphone’ (Lm) line records phonatory movements picked up by a contact microphone placed on the thyroid cartilage and connected by an amplifier to a second electromechanical recorder. The ‘Nose-microphone’ (Nm) line reveals nasal resonances collected in a glass olive placed in a nostril, which communicates via a rubber tube with a microphone connected to a third electromechanical recorder. The ‘Mouth-microphone’ (Bm) line records mouth resonances picked up by a microphone placed in a mouthpiece and connected by an amplifier to an electromechanical recorder. The ‘Mouth-air’ line (Ba) records variations in air pressure exiting the mouth. This pressure is collected at the mouth outlet in a mouthpiece and conducted by a rubber tube into a recording drum covered by an elastic membrane.

Figure 3 shows photos of this acquisition system and Figure 3a is a schematic representation of the system. Figure 3b is a mingogram representing the utterance /purkwatɛʒkasesbatɔ̃/ of the French phrase “Pourquoi t’ai-je cassé ce bâton ?” (Why did I break this stick for you?).

The larynx-microphone (Lm) indicates partial devoicing of /r/ and /ʒ/ due to anticipatory assimilation by the voiceless plosive /k/ and voicing of the otherwise voiceless fricative /s/, due also to backward assimilation by the voiced plosive /b/ (look at the arrows relating to these two sounds). Devoicing is more pronounced for /ʒ/ compared with /r/, as can be observed in the mouth-microphone (Bm), which shows influences from the preceding vowel /u/ extending to the /r/ segment (look at the two vertical lines). Variations in air pressure, recorded by the mouth-air line (Ba), reveal dorsal-uvular flaps characteristic of /r/ in French. An arrow points to the nasality inherent to vowel /ɔ̃/ in the nose-microphone channel. In rare cases, nasal resonances, which could appear in this channel for non-nasal sounds, are rightly interpreted as phonetically irrelevant cranial vibrations.



Figure 3. Photos of made-in-house Mingograph at The Phonetics Institute of Strasbourg as from 1960.

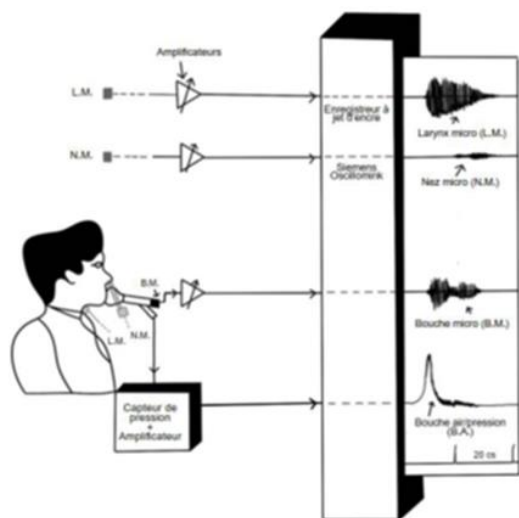


Figure 3a. A schematic representation of the Mingograph

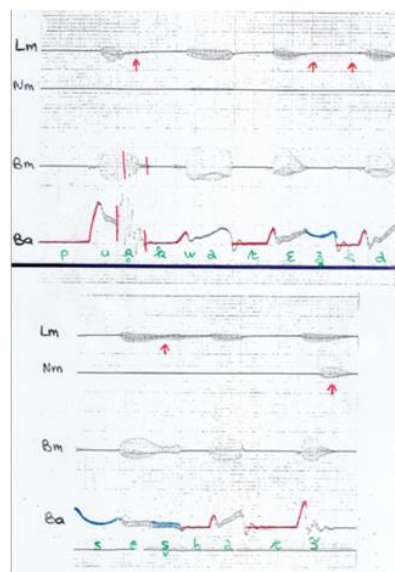


Figure 3b. A mingogram

5 X-ray movies

Combining sounds and images of the vocal tract in the sagittal plane, cineradiographic documents provide essential raw audio-visual information on the geometry of the vocal tract and the movement of the various articulatory structures, synchronised with the acoustic signal (Brock [6], Vaxelaire et al. [7]). These X-ray data, or X-ray movies, offer information, not only on the individual positions of the various speech articulators, but also on the overall shape of the supraglottal cavities responsible, as resonators, for amplifying and modulating the laryngeal tone, i.e. the result of vocal fold activity. X-ray documents were acquired using the well-known principle of the X-ray technique. X-rays, which penetrate matter to a greater or lesser extent depending on its thickness, density, chemical composition, etc., are sent to the organic matter in a uniform beam, perpendicular to the plane to be studied. The X-rays decompose the material they pass through into opaque or clear zones that can be analysed if the X-rays are transformed into visible rays by interposing a screen covered with a fluorescent layer between the material and one's visual inspection. Figure 4 shows a recording session, in 1960, at the Central Radiology Department of the Faculty of Medicine in Strasbourg. The films (50 frames per second, i.e. one image every 20 ms) obtained were projected at full size in order to be hand drawn and obtain measurements. The tracings are drawn from the vertical projection of the film, the projector (Bell & Howell™) being attached to a screw stand, specially designed at The Phonetics Institute of Strasbourg for exploitation of X-ray movies.

Figure 4a illustrates sagittal profiles of a movie of the vocal tract showing (from left to right) the utterance /marɔ̃/ of the French word “marron” (brown). Bilabial contact can be seen clearly for the production of the /m/, whilst opening of the velar-pharyngeal port needs close examination by an expert and painstaking scrutiny in order to be observed, as raising of the velum for the production of the upcoming /a/ is anticipated before release of the bilabial closure. Jaw lowering, coupled with tongue-body lowering and displacement towards the pharyngeal wall to create the necessary constriction at the pharynx for the production of the French vowel /a/, is shown in the second image. The dorsal-uvular contact can be detected in the /R/ configuration and lowering of the velum for the desired opening of the velar-pharyngeal port to produce the nasal vowel /ɔ̃/ can be noticed in the last sagittal profile image.

Figure 4b are hand-drawn tracings of profile views showing the production of sequences /anga/ (left image) in the French phrase “Les ânes **g**ardiens” (The guardian donkeys) and /agna/ (right image) in the French phrase “La **b**ague **n**acrée” (The pearly ring), at normal (sol-

id line) and fast (dotted line) speech rates. Briefly, it can be seen that the apical tongue gesture and the tongue dorsum gesture are rather sequential at a normal speaking rate (solid lines) and occur simultaneously at a fast-speaking rate (dotted lines). These documents tell us that in an intra-articulator context, gestural simultaneity is more difficult to achieve, unless the task is performed in an accelerated manner.

In addition, regarding the gesture of the velum, the lowering or raising of this structure is constrained by the adjacent context. In a vowel context, the velum is “free” to anticipate its lowering movement and to delay its raising movement. This is not the case in a consonantal context, where /g/ seems to be more resistant to nasal influences. Thus, in the sequence /anga/ (left image), the lowering of the velum for /n/ can be largely anticipated in the preceding vowel /a/. However, there is no anticipation in the transition from /g/ to /n/ in the sequence /agna/ (right image). Similarly, raising of the velum may be delayed in the transition from /n/ to /a/, but no such delay is observed in the transition from /n/ to /g/ in the sequence /anga/.

The documents suggest that vowels are less resistant than consonants to nasal influences. However, gestures can be relatively sequential or coproduced, nasalisation can be anticipated or delayed, as long as the realisation of a phonological contrast is not endangered: kinematic strategies and their acoustic consequences must be viable in relation to the constraints imposed by the linguistic code (cf. Sock & Vaxelaire [8]; Vaxelaire [9]).

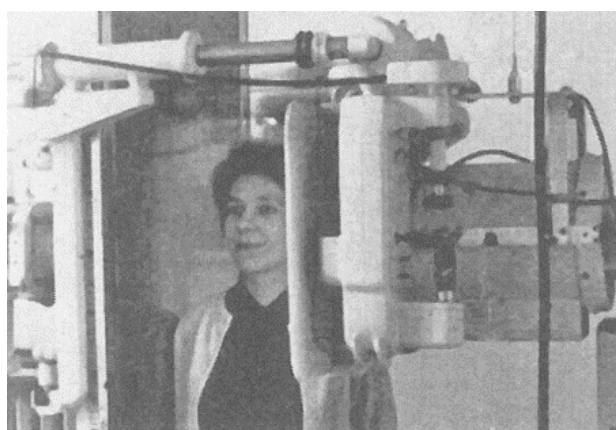


Figure 4. A recording session at the Central Radiology Department of the Faculty of Medicine in Strasbourg (1960).

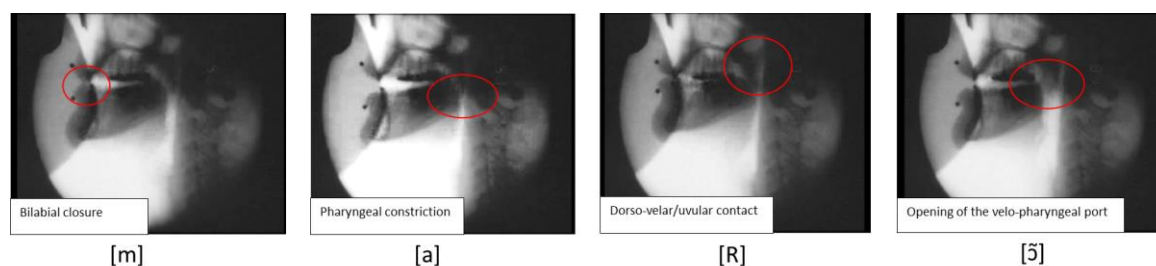


Figure 4a. Sagittal profiles of a movie of the vocal tract showing (from left to right) the utterance /marɔ̃/ of the French word “marron” (brown).

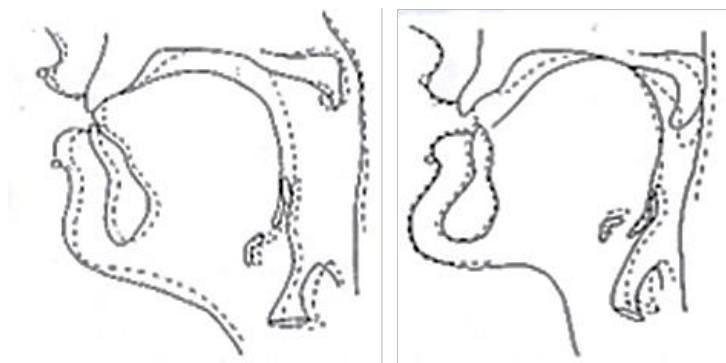


Figure 4b. Hand-drawn tracings of profile views showing the production of sequences /anga/ (left image) and /agna/ (right image) at normal (solid line) and fast (dotted line) speech rates (Vaxelaire [9]).

6 Conclusion

It should be obvious that many heritage scientific documents represent, in general, foundational gems that exemplify primary applications of scientific methods. Such documents indeed provide precious comprehension of the epistemological approach, observational expertise, and experimental procedures adopted by former researchers. Thorough analysis of heritage documents allows one to establish the outline of progression of scientific inquiry from intuition-grounded practices to carefully laid methodologies.

The importance of old phonetic documents resides in the fact that they allow uncovering increasing scientific expertise of different periods. Early researchers in the speech sciences made, at their time, considerable groundbreaking contribution that paved the way for future phonetic investigation. Their findings comprised the principles of reflection, design and creation of adapted instruments, hypothesis formulation, and empirical data, which today are keystones of contemporary phonetic method.

At The Phonetics Institute of Strasbourg, like in many other phonetics laboratories, heritage phonetic instruments and documents influenced various disciplines in the speech sciences, including speech recognition and synthesis, forensic phonetics, clinical linguistics and phonetics, speech therapy.

Today, the institute continues to use palatography, kymography, the Four-line acquisition device, the Sonagraph and many other heritage instruments as teaching tools, but no longer for research. Heritage documents relating to these instruments are of significant didactic value, as they serve as a reference point facilitating the continuity of knowledge, helping to bridge gaps between old phonetic practices and contemporary scientific understanding in the speech sciences.

7 References

- [1] MOSES, E. R.: *A brief history of palatography*. *Quarterly Journal of Speech*, 26(4), pp. 615–625, 1940. <https://doi.org/10.1080/00335634009380603>
- [2] CHAFCOULOFF, M.: *Voir la parole*. *Travaux interdisciplinaires du Laboratoire Parole et Langage*, 23, pp. 23-65, 2004. hal-00285553
- [3] ROUSSELOT, P.-J.: *Principes de phonétique expérimentale*, [Tome 2], H. Welter, Paris, 1897.
- [4] TESTON, B.: *Un système de mesure des paramètres aéro-dynamiques de la parole : le Polyphonomètre, Modèle III*. *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, vol.9, pp. 373-383, 1984.
- [5] ASHBY, P. and M. ASHBY: *The London phonetics training of Masao Kanehiro (1883-1978)*. *Journal of the English Phonetic Society of Japan*, 13, pp. 21-37, 2009.

- [6] BROCK, G.: *Méthode de synchronisation graphique images-son pour l'exploitation de films radiologiques. Présentation de l'appareillage réalisé à l'institut de Phonétique de Strasbourg. Travaux de l'Institut de Phonétique de Strasbourg – TIPS*, 9, pp. 221-232, 1977.
- [7] VAXELAIRE, B., S. PÉLA, G. BROCK, J.-P. ZERLING, F. HIRSCH, F. MARSAC, D. CHARUAU and R. SOCK: *Studies on speech production at the Phonetics Institute of Strasbourg: a historical perspective on the evolution of basic knowledge*. In: Q. LOPES, A. BRAUN and M. ASHBY (dir.): *Proceedings of the Fifth International Workshop on the History of Speech Communication Research*, pp. 87-96, Technische Universität Dresden Press, Dresden, 2022.
- [8] SOCK, R. and B. VAXELAIRE: *How special is speech?*. In: S. FUCHS, H. LOEVENBRUCK, D. PAPE, P. PERRIER (eds.): *Some Aspects of Speech and the Brain*. Peter Lang Internationaler Verlag der Wissenschaften, Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien, pp. 259-294, 2009.
- [9] VAXELAIRE, B.: *Contrôle du velum et nasalité en français. Scolia (Sciences Cognitives, Linguistique et Intelligence Artificielle)*, 10, Publication de l'Université des Sciences Humaines de Strasbourg, pp. 39-60, 1997.

8 Other works consulted in writing this article

- BROCK, G., C. BUFFONI, J. CLARENC, P.-O. GAUMIN and F. HIRSCH: *Les laboratoires de phonétique français dans la première moitié du XXe siècle : instrumentations, corpus et thématiques de recherche*. In: C. DODANE and C. SCHWEITZER (ed.) : *Histoire de la description de la parole : de l'introspection à l'instrumentation*, Paris, Champion, 2021. halshs-03282227
- SIMON, P.: *Les consonnes françaises*, Klincksieck, 1967.
- STRAKA, G.: *Album phonétique*, Québec, Presses de l'université Laval, 1965.

THE PHONETICS LABORATORY OF MONTPELLIER IN EARLY 20TH-CENTURY FRANCE

*Fabrice Hirsch¹, Christelle Dodane^{1,2}, Ivana Didirková¹, Fanny Guitard-Ivent¹,
Laudy Arida¹, Ahamada Kassime¹, Lou Brun¹, Cwiosna Roques¹, Lorelei Berger¹*

¹UMR5267 Praxiling, Université de Montpellier Paul-Valéry & CNRS, France, ²UR7345
CLESTHIA, Université Sorbonne Nouvelle, France
fabrice.hirsch@univ-montp3.fr

Abstract: The Phonetics Laboratory of Montpellier, founded in 1904 by Maurice Grammont, offers a compelling case of early institutionalization of experimental phonetics in France, outside the central academic hubs of Paris and Grenoble. Emerging during a period of technological advancement in speech analysis, the Montpellier laboratory distinguished itself through its dual orientation—both educational and experimental. Equipped with state-of-the-art instruments for the time, it enabled the physiological, articulatory, and acoustic study of spoken language. Maurice Grammont gave the laboratory a unique identity at the crossroads of comparative linguistics, dialectology, and instrumental phonetics. Following Grammont’s retirement, the laboratory entered a period of gradual decline. Scientific research activities were abandoned, and phonetics continued only as a subject taught in literature and language programs. Over time, much of the equipment was lost or scattered, and the laboratory ceased to exist as a formal institutional structure. However, a scientific revival began in the early 2000s—not through the reactivation of the original laboratory, but via the integration of phonetic research into new interdisciplinary research units. This renewal brought updated tools and contemporary themes such as prosody, speech disorders, and corpus phonetics.

1 Introduction

The study of speech has deep historical roots, with auditory-based descriptions of spoken language long preceding the rise of phonetics as a scientific discipline. However, the late 19th and early 20th centuries marked a pivotal shift in how speech was investigated, thanks in large part to the development of new instruments for recording, analyzing, and reproducing sound [1] [2]. This technological turning point gave rise to *experimental phonetics*, a field grounded in the physical and acoustic analysis of speech and distinct from earlier, perception-based traditions.

In France, this methodological renewal was championed by figures such as Abbé Jean-Pierre Rousselot, who founded the first experimental phonetics laboratory at the Collège de France in 1897 [3]. Around him, a generation of scholars helped institutionalize the field through the creation of dedicated research and teaching centers. The *Institut de Phonétique Générale* at the University of Paris and the *Institut de Phonétique de Grenoble* [4] were among the earliest of these, each pursuing distinct yet complementary missions within the rapidly evolving field of phonetics.

While the Paris and Grenoble laboratories have received substantial scholarly attention, other regional initiatives—though less widely known—also played an essential role in shaping the early landscape of phonetics in France. One such example is the Montpellier Phonetics Laboratory, one of the first provincial institutions to engage in systematic research and education in speech sciences. Emerging within a broader national movement to combine pedagogy, instrumentation, and scientific inquiry, the Montpellier laboratory provides a

compelling case study for understanding how phonetic research developed beyond the central academic hubs.

This article explores the history of the Montpellier Phonetics Laboratory, tracing its foundation, institutional context, scientific orientations, and the individuals who contributed to its influence. By reexamining this lesser-known center, the study contributes to a more comprehensive understanding of the early institutionalization of phonetics in France.

2 The Montpellier Phonetics Laboratory

2.1 A pioneering foundation in the early 20th century

The Montpellier Phonetics Laboratory was officially established on April 20, 1904 (Figure 1), by the Ministry of Public Instruction [5]. Its creation was part of a broader effort to structure experimental phonetics as an independent scientific discipline at the crossroads of linguistics, physiology, and acoustic physics. From its inception, the Montpellier laboratory stood out for its experimental and educational focus. It was attached to the Faculty of Arts of the University of Montpellier, where it was based, and enjoyed clear institutional support, notably through ministerial recognition of its status.

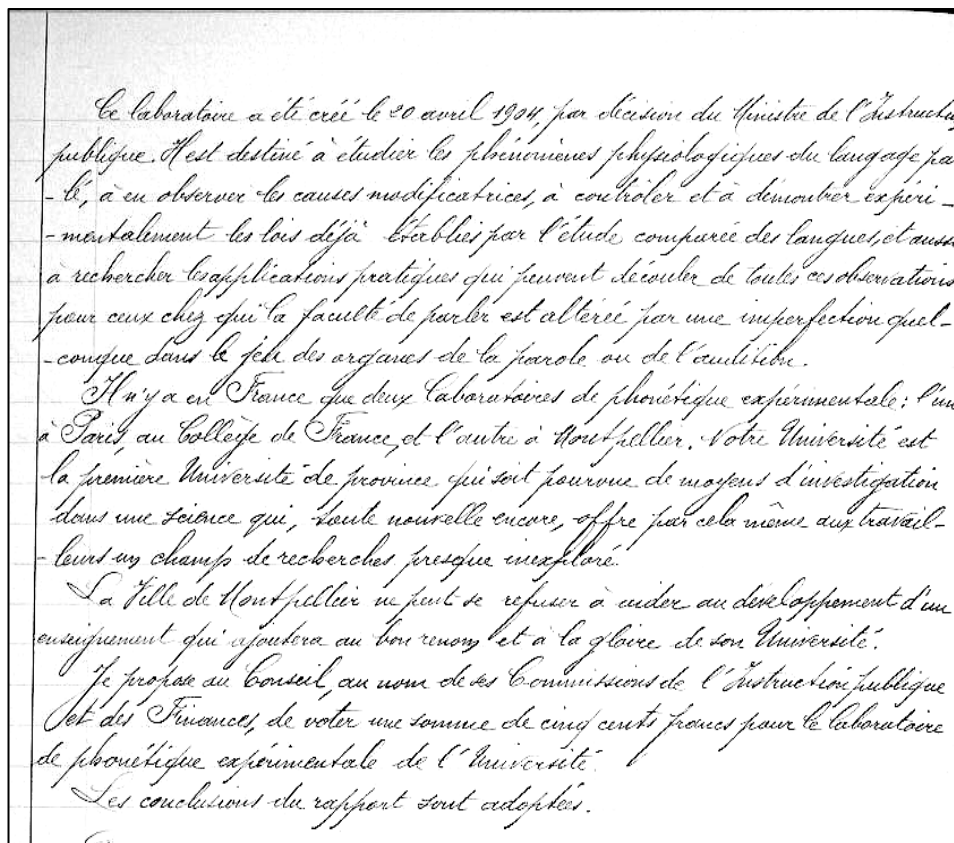


Figure 1. Deliberation of the City of Montpellier dating from 1906 attesting to the establishment of the Montpellier Phonetics Laboratory in 1904. It states in particular that this laboratory is “intended to study the physiological phenomena of spoken language, to observe the causes of its modification, to verify and demonstrate experimentally the laws already established by the comparative study of languages, and also to seek practical applications that may arise from all observations for those whose ability to speak is impaired by an imperfection.” The document indicates that there were only two phonetics laboratories in France at the time: one in Paris and one in Montpellier.

It thus established itself as one of the leading centers for the application of the principles of scientific phonetics at a time when research into speech production, perception, and recording was acquiring new technical tools and methodological protocols.

2.2 Equipment at the cutting edge of experimental practices

From its earliest years, the laboratory was equipped with a range of instruments representative of the technical standards of experimental phonetics at the turn of the century [5]. Clockwork kymographs were used to record variations in air pressure or muscle movement on rotating supports coated with carbon black, and were one of the most iconic tools of early scientific phonetics. They are complemented by recording drums, amplifier levers, metal mouthpieces, nasal olives, and other accessories that allow for detailed exploration of the dynamics of articulatory production. Currently, the laboratory still has a kymograph dating from 1905.

The laboratory also has kymographs (Figure 2), artificial palates (forms 1G & S, models A1, A2, A3, BO, and H2), designed to simulate mouth configurations and facilitate the teaching of articulatory gestures. These objects, designed by Montalbetti, illustrate the alliance between anatomical modeling and articulatory teaching. In addition, there are more specialized devices such as the Atkinson apparatus, dedicated to observing the movements of the soft palate, and Dr. Pescher's spiroscope, used to measure lung airflow during phonation.

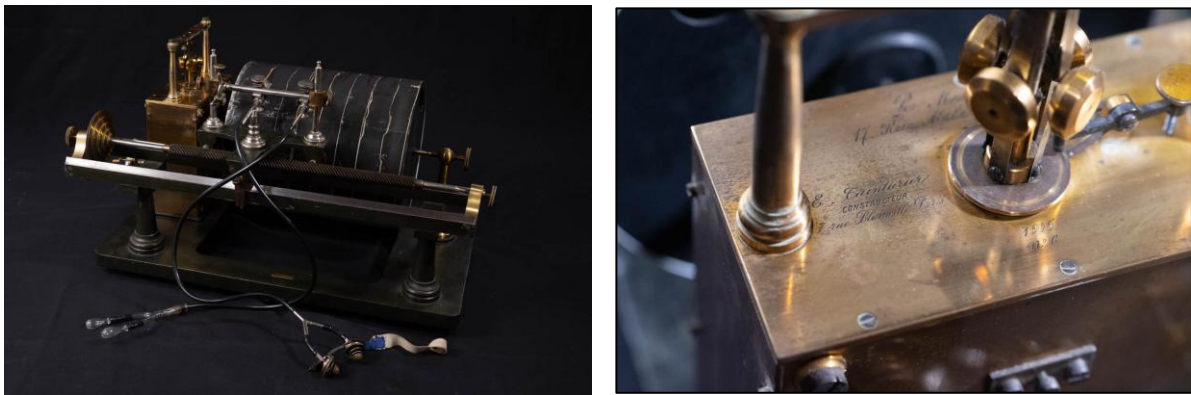


Figure 2. Kymograph dating from 1905, still present at the Université de Montpellier Paul-Valéry.

Another group of instruments is designed to study the acoustic dimension of sounds. These include tuning forks—simple, slider, and electro-tuning forks—used to produce and calibrate known frequencies, which are essential for acoustic demonstrations and auditory perception exercises. These tuning forks are mentioned in equipment inventories from the laboratory's early decades and testify to an early interest in the physical properties of sound [6]. The use of these instruments reflects the laboratory's ambition to be part of a multidisciplinary tradition, combining anatomical analysis, articulatory dynamics, and acoustics.

2.3 The structuring role of Maurice Grammont

The growth and scientific influence of the laboratory are closely linked to the figure of Maurice Grammont, a linguist and phonetician with a classical education, who played a central role in structuring the field of instrumental phonetics in France. Founder of the institution, he was both its first director and the guarantor of its scientific orientation. His work, at the crossroads of historical linguistics, experimental phonetics, and language teaching, drew on the laboratory's instruments, which were used for both research and training purposes. Grammont actively participated in the integration of experimental methods into the teaching of phonetics, convinced that instrumental observation made it possible to overcome the limitations of subjective perception.

He attracted several foreign researchers to Montpellier to train in experimental methods. Among them, Ebenezer Suddard is a prime example: after training in Montpellier, he went on to introduce scientific phonetics to Perth, Australia [7]. This international influence helped establish the Montpellier laboratory's reputation as a center for training and dissemination of experimental practices.

2.4 A gradual decline after Grammont's departure

The retirement of Maurice Grammont between the two world wars marked a turning point in the institutional history of the laboratory. His chair was entrusted to a linguist who was not involved in the field of phonetics, which led to a gradual disengagement of the Faculty of Arts from the discipline. Phonetics continued to be taught, but was entrusted to lecturers, assistants and then assistant professors, who did not carry out any research.

This reorientation led to a narrowing of the laboratory's scientific role, which essentially functioned as a teaching support facility. The existing equipment was used for teaching purposes, but the lack of a maintenance and renewal policy hampered its development. Nevertheless, some instruments continued to be replaced or modernized, demonstrating a form of material persistence, even in the absence of a real structuring scientific project.

2.5 Instrumentation maintained but scattered

An inventory carried out in 1956 [6] provides an overview of the equipment still in use in the mid-20th century. In addition to the production devices mentioned above, there were cathode ray oscilloscopes, a Thomson microphone, Montalbetti cylindrical resonators for studying resonant frequencies, a Reichert microscope for detailed analysis of tracings, and several sound recording and reproduction devices: tape recorders, Ducretet-Thomson electrophones, and Télétronic tape recorders. The collection is rounded out by a series of Montalbetti tongue guides, designed for corrective pronunciation training.

However, this wealth of material has eroded over the decades, mainly due to the laboratory's multiple moves within the university, which have resulted in the loss or deterioration of a large number of historical instruments. While some tools were kept in storage or collections, many disappeared or became inoperable due to lack of maintenance or technical documentation.

Although the phonetics laboratory itself eventually ceased to exist as a dedicated institutional structure, phonetics continued to be taught—primarily within specific academic programs such as degrees in literature and language.

2.6 Towards a scientific revival in the 2000s

In the 1980s, the status of assistant professor was abolished by ministerial reform. The last holders of this status were automatically promoted to associate professors, without their teaching duties necessarily being accompanied by an explicit research mission. As a result, phonetics remained present in the curriculum, but the former laboratory no longer functioned as an institutional or scientific entity.

A genuine scientific revival did not begin until the early 2000s, with a new wave of recruitment. The teacher-researchers who succeeded the former incumbents were specialists in phonetics, often trained in research and actively involved in collective projects and national or international networks. This generational transition enabled the resumption of phonetic research at the university, now integrated into interdisciplinary research units rather than a dedicated phonetics laboratory. It was supported by contemporary research themes (prosody, inter-speaker variability, speech disorders, corpus phonetics, etc.).

The history of the Montpellier phonetics laboratory illustrates the tension between ambitious foundations, institutional decline, and partial reinvention. It was initially a cutting-edge center for phonetic experimentation, in keeping with the spirit of the emerging speech sciences. The withdrawal of the discipline from the local university scene then led to a loss of its institutional and scientific anchoring. However, the persistence of a material base, educational continuity, and the quality of training helped preserve a core of phonetic expertise.

Since the beginning of the 21st century, phonetic research has found new momentum in Montpellier, not through the revival of the former laboratory, but through its integration into

existing research structures. At the crossroads of speech technologies, experimental studies, and the humanities, it is now part of a scientific tradition in constant transformation—faithful to the legacy of Maurice Grammont, while embracing the contemporary challenges of phonetic inquiry.

3 A word about Maurice Grammont and the laboratory's initial research

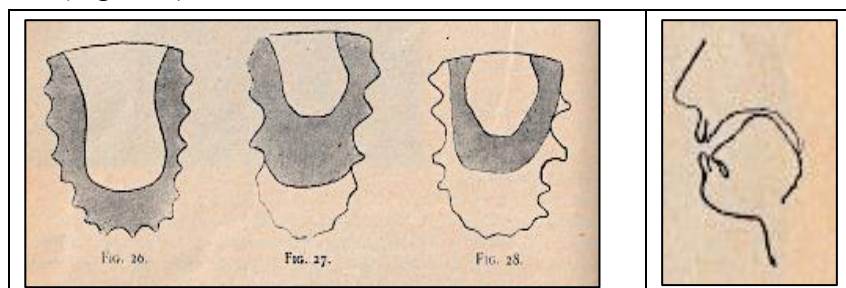
Maurice Grammont, born on April 15, 1866, in Damprichard (France), was a major figure in the history of language sciences in France.

After obtaining his agrégation in grammar in Paris in 1890, he began his academic career in 1893 as a lecturer at the Faculty of Arts in Dijon, before being appointed two years later to Montpellier, where he taught comparative grammar and philology.

It was at the Sorbonne on March 13, 1896, that he defended his doctoral thesis in literature entitled *La dissimilation consonantique dans les langues indo-européennes et dans les langues romanes* (Consonant Dissimilation in Indo-European and Romance Languages) [9]. This work marked a milestone in the study of phonetic regularities: Grammont formalized some twenty general phonetic laws, which he considered valid *beyond languages* as universal manifestations of human language. These laws were not simply empirical regularities: they were interpreted as deriving directly from the physiological structure of the speech organs and the mechanism of articulation. In a famous passage, Grammont defines them as “the laws that govern the evolution of sounds, imposing such modifications in a constant and absolute manner, whenever no particular circumstance prevents them from acting” [9, p. 111]. This approach, which is both systematic and naturalistic, anticipates a biophysiological conception of historical phonetics, which will be further developed in the work of structural phonology and articulatory phonetics.

In 1901, Maurice Grammont became professor of comparative grammar, then, in 1906, Professor of linguistics at the Faculty of Arts in Montpellier. It was in this context that he founded the Montpellier Laboratory of Experimental Phonetics on April 20, 1904. Grammont's approach remained faithful to his conception of language as a physiological phenomenon that could be measured and analyzed through experimentation.

His major work, the *Traité de phonétique*, published in 1933 [10], is a monumental synthesis of the descriptive phonetics of his time. The book is divided into three main parts: phonology, evolutionary phonetics, and impressive phonetics. In the first part, Grammont presents a typology of phonemes, which he classifies according to their articulatory and combinatory characteristics: occlusives, spirants, moistened consonants, oral and nasal vowels, etc. But the originality of the *Traité* lies in the close link it establishes between the articulatory description and the acoustic description of sounds. This integrative approach, at a time when acoustic analysis tools were still limited, demonstrates his desire to establish a comprehensive phonetics, combining anatomy, articulatory dynamics, auditory perception, and physical analysis of sound (Figure 3).



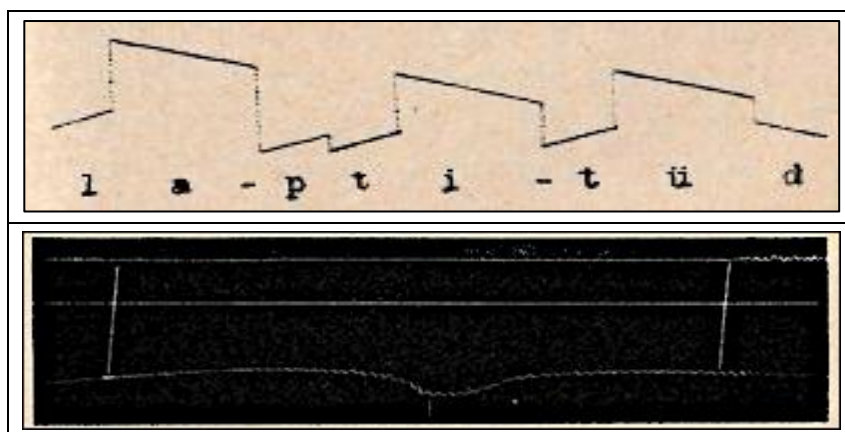


Figure 3. examples of figures in the *Traité de Phonétique*: top left, palatographic representations (here of [t] in French, [td] in pre-palatal English and [t] in Dravidian languages [10, p. 48]. Top right: sagittal section of dorsal [r] [10, p. 73]. In the middle, a representation of the articulatory tension phenomena associated with the different syllables in the statement “l'aptitude” [10, p. 101]. At the bottom, three kymograph drums, with the mouth orifice on the upper line, the nasal orifice on the middle line and the larynx on the lower line [10, p. 62]. These figures are numerous throughout the book (179 in all).

Maurice Grammont thus occupies a unique place in the history of language sciences: as a philologist, phonetician, comparative linguist, and educator, he helped establish the methodological foundations of French scientific phonetics. His influence extended far beyond Montpellier: he trained several generations of teachers and researchers, some of whom went on to continue his work internationally.

4 Scientific influences on Maurice Grammont and the Montpellier phonetics laboratory

Grammont was not content to simply reproduce existing models: he gave the laboratory an original orientation, at the crossroads of comparative linguistics, experimental phonetics and dialectology. Through him, the Montpellier laboratory became a place of technical experimentation, but also of theoretical formalization, in a provincial university setting that favored a certain intellectual autonomy.

4.1 An academic career marked by the great names of linguistics

Maurice Grammont's initial training was part of a network of prestigious influences. He followed the teachings of Michel Bréal, Ferdinand de Saussure, Louis Havet, Victor Henry, Henri d'Arbois de Jubainville and Johannes Schmidt, and studied comparative linguistics, Sanskrit, Gothic, Lithuanian and Old High German. He also attended classes given by Abbé Jean-Pierre Rousselot, from whom he adopted certain principles relating to the instrumental study of speech. In his thesis, defended in 1896, Grammont paid tribute to several of these intellectual figures, indicating the diversity of the scientific traditions that influenced him [8].

This plurality of sources of inspiration is reflected in his own method, as he himself puts it: “turning to a new master as soon as he thought he had grasped the method of the previous one”. The phonetics laboratory he founded in Montpellier was therefore conceived not as a replica of an existing model, but as a space for implementing this theoretical plurality.

4.2 A laboratory at the crossroads of phonetics and comparative linguistics

From the outset, the Montpellier laboratory was conceived as a place dedicated to the physical observation of speech, with both pedagogical and scientific aims. It was soon equipped with state-of-the-art equipment for the time: tuning forks, kymographs, sound recording and visualization devices. This orientation towards instrumentation was in line with Rousselot's

methodological heritage, but Grammont's use of it was part of a different approach: the aim was not simply to produce tracings, but to derive general laws on the evolution of sounds.

It was from this perspective that he developed his work on dissimilation, phonetic laws and the universal tendencies of language. The laboratory thus became the material extension of linguistic research aimed at explaining phenomena, rather than simply describing them.

4.3 Anchored in the French School of Linguistics

Grammont helped define the contours of the French School of Linguistics, which he mentioned as early as 1902. The school's distinctive method focused on the study of the general characteristics of languages, their evolutionary tendencies and universal phenomena. In Montpellier, these principles found concrete application, notably in the laboratory's activities and the teaching courses it organized, including those for foreigners - the forerunner of today's Institut Universitaire d'Enseignement du Français Langue Etrangère (IEFE) [8].

Through his participation in academic networks (*Société de Linguistique de Paris*, *Revue des Langues Romanes*), Grammont asserts the laboratory's role as a center of linguistic reflection integrated into national scientific life. His aim was not to build a school around himself, but to make Montpellier an active center for the production and dissemination of linguistic knowledge.

4.4 Measured but real international influence

The laboratory's influence is not confined to the French scene. Several foreign researchers passed through, including Ebenezer Suddard [7], who would later form the first phonetics laboratory in Australia. This circulation, however limited, testifies to the recognition of the expertise developed at Montpellier under Grammont's direction. Far from the Parisian centers or the networks of the *Association Phonétique Internationale*, the laboratory nonetheless remained a place of experimentation recognized for its scientific rigor. The attention paid to instruments, articulatory correction, phonetic pedagogy and the physiology of speech is part of its own tradition, adapted to the orientations of its founder.

5 Conclusion

The Montpellier Phonetics Laboratory stands as a compelling case of early 20th-century scientific innovation in a provincial academic setting. Founded in 1904 and driven by the vision of Maurice Grammont, it embodied a distinctive approach to phonetics—one that blended experimental rigor, physiological analysis, and linguistic theorization. Grammont's training under leading figures such as Rousselot and Saussure, combined with his philological grounding, enabled him to shape the laboratory as both a research center and a pedagogical platform.

Unlike the central institutions in Paris or Grenoble, Montpellier pursued its own scientific trajectory, marked by its commitment to articulatory precision, phonetic instrumentation, and the formulation of general laws of sound evolution. This approach positioned the laboratory as a foundational yet autonomous contributor to the development of phonetics in France. Though its international influence remained modest, it played a critical role in training practitioners and disseminating experimental techniques—most notably through figures such as Ebenezer Suddard.

The laboratory's decline in the mid-20th century, following Grammont's retirement, reflects broader institutional shifts in French academia. Yet the persistence of its equipment, teaching legacy, and disciplinary know-how ensured a continuity of phonetic practices that later facilitated its scientific revival in the 2000s. Today, the laboratory contributes to contemporary research in speech sciences, reaffirming its place in a tradition that began over a century ago.

By examining the Montpellier Phonetics Laboratory through its material history, institutional developments, and intellectual foundations, this study highlights the value of decentralizing the historiography of phonetics. It reveals how regional laboratories, even when operating outside dominant academic centers, played a crucial role in shaping the experimental, pedagogical, and theoretical frameworks of modern phonetics.

References

- [1] BOË, L.-J., C.-E. VILAIN: *Un siècle de phonétique expérimentale : fondation et éléments de développement*, Lyon : ENS Éditions, 340 p., 2010.
- [2] DODANE, C., C. SCHWEITZER: *Histoire de la description de la parole : de l'introspection à l'instrumentation*, Honoré Champion, 406 p., 2021, 9782745355959.
- [3] HOLZER V.: *L'abbé Rousselot, inventeur de la phonétique expérimentale*. Transversalités, Paris : Institut Catholique de Paris, vol. 171, 214 p., 2024.
- [4] BOË, L.-J., J.-F. BONNOT: Création et émergence de la phonétique expérimentale en France: de l'abbé Rousselot à Théodore Rosset. In *Un siècle de phonétique expérimentale : fondation et éléments de développement*, 103-128, 2010.
- [5] BROCK G., C. BUFFONI, J. CLARENC, P.-O. GAUMIN, F. HIRSCH: Les laboratoires de phonétique français dans la première moitié du XXe siècle : instrumentations, corpus et thématiques de recherche. In *Histoire de la description de la parole : de l'introspection à l'instrumentation*, 301-317, 2021.
- [6] PERROT J.: Laboratoire de phonétique de l'Université de Montpellier (Laboratoire Maurice Grammont). In *Instituts de phonétique et archives phonographiques*, commission d'enquête linguistique, 229-232, 1955.
- [7] ASHBY M.: Australia's first phonetics laboratory, 1913: its founder and its context. In *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 2019.
- [8] DODANE C., F. HIRSCH: Maurice Grammont, une pensée moderne au carrefour de plusieurs disciplines linguistiques. In *Linguistes montpellierains du 20ème siècle : unité et diversité, apports et enjeux*, @racné, 105-120, 2025.
- [9] GRAMMONT, M.: *La dissimilation consonantique dans les langues indo-européennes et dans les langues romanes*, thèse de doctorat, 215 p., 1895.
- [10] GRAMMONT, M.: *Traité de phonétique*. Librairie Delagrave, 480 p., 1933.