

4. Experimental phonetics and phonology

BIOGRAPHICAL NOTES

Ivana Didirková is Associate Professor of Phonetics and Phonology at the Université Paris 8 Vincennes – Saint-Denis. In her research, she is particularly interested in stuttered speech production, articulatory phonetics, and disfluency in speech and discourse.

<https://orcid.org/0000-0001-8107-2361>

Anne Catherine Simon is full Professor of French Linguistics at the University of Louvain. Her research interests cover prosody and spoken syntax within the frame of corpus and experimental linguistics, with a special emphasis on discourse units, regional prosody, and speaking styles. <https://orcid.org/0000-0002-7202-842X>

Abstract

Experimental phonetics and phonology study speech production, perception, and language sound patterns. When perceiving continuous speech, the human ear makes categorisations that sometimes contradict acoustic cues in the speech signal. The role of experimental phonetics and phonology is to model the constraints linked to human speech production and perception.

The two disciplines are tightly linked to the development of multiple devices allowing for speech signal observation. With new technologies, most instruments and techniques developed over a century are now computerised, allowing for detailed speech signal visualisation, deducing the underlying articulatory behaviours and understanding the subtle elements essential in perceiving and interpreting acoustic events. Articulatory movements can now be observed directly, using machines developed for this purpose or diverting medical devices.

Experimental methods are also helpful for phonological questions, such as describing and categorising under-endowed languages. They are also crucial in studying many questions related to discourse. Despite this fact, one of the critical issues in the two disciplines concerns using the most natural data possible, allowing connections between the descriptive and explanatory approaches since laboratory data are controlled and elicited.

1. Introduction and Definitions

Experimental phonetics and phonology aim to understand the processes of speech production and perception. As stated by Ohala, experimentation is an attitude: ‘first, a keen awareness that the world is not necessarily as it may seem, i.e., that our sense-impressions and therefore the opinions and beliefs based on them may be faulty, and, second, the willingness to actively do something to compensate for or correct these potential errors by making observations under carefully controlled conditions.’ (Ohala, 1987, p. 207)

The experimental approach is characterised by the effort to control for factors that could potentially confound their effects on the phenomenon one seeks to explain, model, or predict (Xu, 2011, p. 88). An experimental study is based on a hypothesis (e.g., the higher-pitched a voice is, the more female-like it is perceived) that can be tested by manipulating variables (e.g., the mean pitch of the voice, but also intonation contours or the timbre). The testability of the hypothesis depends on the identification of the variables to be observed and the reliability of the measurements of these variables (Barbosa, 2012). The reliability of the measurements often requires the use of instrumental acoustic analysis (Hayward, 2000, p. 5). However, accurate acoustic analysis is not sufficient if the hypotheses and variables are not properly identified.

That is the reason why experimental phonology cannot be reduced to the biophysical aspects of speech sounds and their computerised analysis. Hypothesis testing crucially relies on related fields of knowledge, depending on the research questions: the explanation of phonological phenomena of contrast and categorisation requires a cognitive dimension (Demolin, 2012); psychological models are needed when investigating language acquisition and development; sociolinguistics allows for studying how we categorize people based on voice features, etc.

2. Historical Perspectives

As in other disciplines, the development of new techniques in the late 19th and early 20th centuries opened new horizons in speech sciences. The apparition and access to new recording machines permitted acquisitions and studies of higher amounts of data and were supposed to allow for more objectivity since a recording can be played back, unlike direct observations made before. Traditionally, it is considered that experimental phonetics started with Rousselot (1846-1924), who introduced the term in 1889 (Boë & Vilain, 2010). His *Principes de phonétique expérimentale* (1924) provides the basis for the discipline,

explaining the necessity of a ‘control’ over subjective observations made by early descriptive phoneticians. Rousselot’s work describes acoustic and physiological parameters of human speech and provides valuable information on ‘natural’ and ‘artificial’ (Rousselot, 1924, p. 4) means of observation and experimentation. The author explains techniques such as sound recording machines (chronographs, phonautographs, phonograph, gramophone) or devices allowing for the observation of articulatory movements (artificial palate, kymograph, myograph, pneumograph, labial, lingual, velar, and laryngeal ‘exploration’). Importantly, Rousselot’s conception of phonetics was multidisciplinary: phonetics included linguistics, acoustics, and physiology. However, while Rousselot is considered the father of experimental phonetics (Hayward, 2000, p.35), the question related to speech studies was raised long before the 20th century. For instance, in the 18th century, Darwin (1731-1802) led the first instrumental phonetic study on a live speaker to describe vowels’ place of articulation (Ohala, 2010).

Another seminal text was published three years after Rousselot introduced experimental phonetics. Scripture’s ‘Elements of Experimental Phonetics’ (1902) introduce what will later be considered as the three main branches of phonetics: speech acoustics (called ‘curves of speech’), speech perception, and speech production, and provides essential information on them.

Experimental phonology is often seen as later than phonetics: Demolin (2012) mentions a nearly one-century delay between Rousselot and experimental or laboratory phonology apparition. Nevertheless, as for experimental phonetics, it is important to specify that, as stated by Ohala (1987), although they probably did not think of themselves as ‘experimental phonologists,’ linguists and anthropologists such as Sapir (1884-1939), Greenberg (1915-2001), or Brown (1925-1997) did use experimental methods in some of their works.

However, it is Ohala and Jaeger who contributed to the development of the discipline. In their *Experimental phonology* (Ohala & Jaeger, 1986, p. 2), they advocate the use of experimental methods in phonology not to create new knowledge, but instead to refine the existing knowledge and recall that experimental methods do not necessarily imply heavy equipment. The authors insist on two facts. First, many other sciences switched to experimental methods that allow for new developments in their discipline, albeit imperfect. Second, theories, often opposed to experiments, would benefit from experiments and vice versa: indeed, an experiment is often based on a theory (or model); otherwise, it would not make sense. Inversely, theories need experiments to be confirmed – or not –and refined if necessary.

3. Critical Issues and Topics

3.1 Variability and complexity

For several decades, experimental phonetics and phonology have pursued the elaboration of a feature theory and sought generalised and straightforward ways to describe phonemic invariants in languages through articulatory or acoustic theories (Ladefoged, 2004).

Invariance was a central concern until Labov's (1972) analysis of sociolinguistic issues highlighted the social variability in sound production and perception.

Today, researchers value the intrinsic and extrinsic variability of speech, including in experimental settings where variables must be controlled. Intrinsic, or structural, variability is due to different parameters (e.g., frequency of a phoneme, syllabic, or prosodic structure) that influence the realisation of phonemes (Meunier, 2005, p. 367). On the other hand, sources of extrinsic variation are numerous (e.g., speaker characteristics, situational context, dialect) and must be considered in production and perception experiments (Thomas, 2002). The result is a complex and multifactorial attempt to model intrinsic and extrinsic variation that has become central in many experimental studies (Kim & Tilsen, 2022; Lawson & Stuart-Smith, 2021).

3.2 Corpora and stylistic diversity

Corpus studies and experimental studies sometimes seem irreconcilable: the term 'laboratory phonology' evokes the use of elicited speech that can be far from everyday speech. However, researchers seek to produce explanatory models applicable to speech in everyday life (whether for automatic recognition, speech synthesis, speech therapy or teaching purposes). Authentic speech corpora (continuous speech, dialogue) are necessary for the results to have ecological validity (Xu, 2012, p. 91). Corpus data come from natural contexts and make it possible to study issues related to stylistic diversity (Wagner et al., 2015). Using such data in experimental studies combines the advantages of both corpus and experimental methods, which are complementary. In some cases, corpus and experiment are on the same footing and the study is both observational and experimental; most of the time, corpora serve as databases from which fitting examples are retrieved (Gilquin & Gries, 2009, p. 14).

3.3 Multimodality

A last critical topic we briefly dwell on is the need for adopting a multimodal perspective. Multimodality first relates to the mode (spoken, written or gestural) of speech production and the interaction of auditory, visual, and gestural modes when processing speech. Scholars

study the influence of written stimuli on phoneme categorisation (Dufour et al., 2022), the interplay between gestures and prosody in speech production and interpretation (de Ruiter et al., 2012, Prieto et al., 2015), or the phonology and prosody in sign languages (Brentari, 2019). A second meaning, more abstract, relates to the fact that speech can be conceived as multimodal in itself: the interplay of speech rate, breathing, register and pausing, alongside the production of accentuation, rhythm or intonation patterns, creates a complex system where the different prosodic ‘modes’ are partly autonomous and interplay with each other (Perlman & Benitez, 2008; Simon & Auchlin, 2001).

4. Current Contributions and Research

4.1 From typology to variation

The field of experimental phonetics and phonology is incredibly rich and covers various topics. Initially, descriptive methods used by the two disciplines were meant to understand the basic functioning of the vocal tract, such as articulatory descriptions of different speech sounds or laryngeal behaviour, both in typical and in impaired speech. Since then, the topics covered still include phonological systems descriptions but also focus on other research questions.

4.1.1 Language typology and description

Despite decades of studies aiming at speech sound description, there still are languages containing phonetic or phonological structures that are yet to be understood. Recently, a group of researchers published a book on laryngeal features in Native American languages (Avelino et al., 2016), containing phonetic and phonological studies on parameters such as tones, vowel laryngealisation, temporal coordination of glottalic gestures or consonant-tone interactions. Authors use a large scale of phonetic analyses, including acoustic or electroglottographic studies. Other researchers are interested in African languages, investigating their rhythmic patterns (Sedhu, 2015), intonation (Zerbian & Barnard, 2008), articulatory gestures (Demolin & Chabiron, 2013), and others. South American languages also gave rise to a book published a decade ago (Campbell & Grondona, 2012) and studying, among other parameters, their phonetics, and phonology. Finally, Asian language specificities are also widely studied, often regarding their tone inventories (Brunelle & Kirby, 2016; Michaud, 2012). Interestingly, besides these less endowed languages, there is still a need for a more detailed understanding of occidental languages and some of their phonetic and

phonological features. For instance, Signorello et al. (2017) recently investigated the aerodynamic characteristics of Belgian French fricative consonants by measuring subglottal and intraoral air pressure. Rhoticity and its articulatory and acoustic parameters are also a subject of interest in different languages such as English (King & Ferragne, 2020), Greek (Nicolaidis & Baltazani, 2014), Slovak (Pouplier & Benus, 2009), and others, or in different configurations such as articulation problems (Van Lieshout et al., 2008).

4.1.2 Variation

While we do have detailed information on phonetic properties of many speech sounds, syllables, coordination, and phonological systems of languages, variation in these parameters is not scarce. On the contrary, many studies try to disentangle essential phonetic and phonological features of a sound/structure and features that can vary without preventing the correct identification.

As mentioned in the previous section, variation in speech and language is caused by many factors and is considered normal. Current research on variation in speech production and perception focuses on a number of these factors.

One of the active fields is inter-, and intraspeaker variation or, in other words, the way speech parameters vary not only between two speakers but also how one individual makes different speech production characteristics vary, depending – or not – on other variables. Several studies are interested in the variability of articulatory movements over time, depending on the speaker's age and other factors (Grigos, 2009; Jacewicz et al., 2010; Tomaschek et al., 2021). One of such factors, and another important question that is still being addressed, is geographic variation. For illustration purposes, in English, the wide variety of accents gives rise to studies on speech production and perception of different English spoken varieties in comparison (or not) with comparable speech-related structures in other languages (Cebrian, 2021; Przewozny-Desriaux et al., 2020). Intraspeaker variation can also be due to differences in the communication setting. It is long known that formal situations require a more careful speech than casual discussions. However, many other, more detailed analyses are still led on this topic, studying the relationship between prepared and spontaneous speech, media or non-media speech, interaction degree, and others (Brognaux & Drugman, 2014; Goldman et al., 2014), both in speech perception and speech production.

4.1.3 Pathological speech

Another important field, clinical phonetics and phonology use both perception and production paradigms to assess specificities in speech and language disorders, hearing impairments, and other disorders affecting speech and language. This subfield implies collaborations with speech and language pathologists, neurologists, otorhinolaryngologists, and other specialists. When focusing on speech and language production, studies in the clinical field often use different vocal tract visualisation tools. One of their most important aims is to identify similarities and differences between non-pathological and pathological speech behaviour, helping thus to understand not only the disorder but also the typical function of speech, improving different speech and language-related models and theories. Perception studies, on the other hand, focus on how disordered speech and voice can influence communication and how disorders are perceived. Studies in this field are interested in disorders such as apraxia of speech (Overby & Caspari, 2015), dysarthria (Lehner et al., 2022), stuttering (Didirková et al., 2021), or different voice disorders (Awan, 2008), and prosody impairments (Wells & Whiteside, 2008).

4.2 From segments to discourse

While the aforementioned studies mainly focus on phonetic and phonological parameters, scientists also use the two experimental sciences to link the speech level and a higher discourse level to help understand which voice and speech features contribute to the global sense of the message delivered by the speaker.

4.2.1 Prosody

Prosodic (suprasegmental) features are one of the determinant parts of spoken language and its processing. They provide information on lexical, syntactic, and discourse levels, the speaker, their emotional state, the language variety, and the speech style used, and other parameters. Literature on this subject is rich and uses many experimental methods investigating prosodic features of various spoken language structures. At the syntactic level, studies have explored the syntax-prosody interface in many languages (Dogil et al., 2002; Georgiafentis & Sfakianaki, 2004; Gras & Elvira-García, 2021; Hsiao, 2020; Ma & Zhuang, 2022; Michelas & D’Imperio, 2012). The same is true for the discourse-prosody interface, for which studies investigate the role of prosody in the interpretation of the discourse relation (i.e., consequence, specification, topic shift, and others, see Fischer, 2006) it conveys (Didirková et al., 2019; Gonen et al., 2015; Van Praet & O’Grady, 2018), or its role in discourse structuration (Tyler, 2013). Prosody is also investigated in children and/or in

various speech and hearing-related impairments (for example, in inferring discourse prominence in cochlear-implant users (Huang et al., 2017), in investigating syntax-prosody interface in children with developmental language disorder (Caccia & Lorusso, 2019), in ageing (Steinhauer et al., 2010).

4.2.2 Disfluency

Speech production is a highly demanding process requiring the completion of several steps before speech sounds are actually released (for more information, see for example Guenther, 2016; Levelt, 1989; Shattuck-Hufnagel, 2019). This complexity often leads to hesitation phenomena, especially in unprepared speech, as the speaker has to prepare and produce their speech in real time. What is commonly experienced as a lack of words, a hesitation in the formulation, the time one takes to think while producing ‘uh’ and other ‘er, well’ is commonly called disfluency and gives rise to a large number of studies in linguistics, discourse analysis, but also in speech sciences. Disfluencies are investigated in typical speech in terms of prosody (Wu et al., 2022), speaking style (Moniz et al., 2014), speech recognition (Stouten et al., 2006), discourse (Merlo & Mansur, 2004), speech production (Anansiripinyo & Onsuwan, 2019), or speech perception (Warner et al., 2022).

5. Main Research Methods

Experimental phonetics and phonology are mainly based on instruments allowing for more or less direct observations of the parameters considered. This section will list some of the tools used in present-day studies.

5.1 Speech perception studies

Speech perception is an essential part of speech since the act of speaking is generally realised with a communication purpose. This includes not only the message transmission per se, but also the speaker’s age, emotional state, speaking style, and other important communication cues. Several experimental tasks are found in the literature regarding perception (Table 1).

5.1.1 Listening studies

The simplest method for a perception study consists of making participants listen to several stimuli and answer the related questions (see Table 1). Many platforms are available for such studies (e.g. PsychoPy, Google Form, LimeSurvey). Some of these platforms allow for measuring reaction time or performing actions like sensorimotor synchronisation with sound

stimuli by finger tapping along with looped spoken phrases (Rathcke et al., 2021). Another essential advantage is the possibility to recruit participants online, which allows for a more considerable number of participants.

Table 1: Main experimental tasks in listening studies

Categorization and judgment tasks in listening studies (see Thomas, 2002: 135; Gilquin & Gries, 2014: 5)	(1) judging whether two stimuli sound the same or different; (2) judging which two stimuli out of three or more are alike (ABX paradigm); (3) sorting stimuli, freely or into categories; (4) gauging the acceptability, grammaticality, naturalness... of speech stimuli (categorically or on a Likert-scale)
Identification or transcription tasks	(1) identifying whether or not a stimulus contains a typical feature; (2) transcribing sound features (prominence, focus, prosodic boundary) using e.g. the rapid prosodic transcription (RPT) protocol (Cole et al., 2017)
Anticipation	Gating

More often, however, perception studies include the use of other instruments. The gating paradigm (Grosjean, 1980) assesses the hearer's anticipation of the sequence to come. In other words, the hearer is repeatedly presented with a spoken stimulus, and the duration of the stimulus (from its onset) is increased at each successive presentation. The presentation time, called gate, varies according to the stimulus type (syllable, word, sentence), and the hearer is asked to indicate their expectations in terms of the sequence to come (typically one of several propositions), but also their confidence (Moradi et al., 2014).

5.1.2 Stimulus preparation

Speech synthesis is widely used in speech perception studies, in order to modify natural stimuli to test one or several particular parameters (e.g. sound duration, formants) while controlling for others (see, for example, Didirková et al., 2019). Using the prosody-

transplantation paradigm allows for transplanting the prosody of one speaker to a sentence produced by another one speaker, mixing segmental parameters from one variety to suprasegmental parameters from another variety (Boula & Vieru-Dimulescu, 2006). As for studying the perception of prosody, researchers frequently produce delexicalised stimuli, where segmental information is altered, so that the hearer concentrates on prosodic parameters only without being influenced by linguistic content. Such content masked speech is produced by using meaningless words or sentences (Bänziger & Scherer, 2005), by low-pass or stop-band filtering (Goldman et al., 2014), by different kinds of resynthesis trying to keep the resulting stimuli as natural as possible, that is avoiding humming (phoneme substitution Simon & Christodoulides, 2016: 87; inverse filtered glottal flow Vainio et al., 2009).

Magnetoencephalography (MEG), electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and eye-tracking are sometimes used to assess perception from a physiological viewpoint. More precisely, they describe the human brain's response to speech stimuli (Alday, 2019; Redcay et al., 2008; Wang et al., 2020).

Speech production studies

Multiple tools and measurements help to objectify speech production parameters. In general, phoneticians and phonologists are interested in breath, voice, and speech characteristics (Table 2).

5.1.3 Acoustic analyses

One of the most obvious instrumentations is a simple voice recorder or any other device (computer, smartphone) equipped with a sound card, a microphone, and software allowing voice recording. Audio data issued from such recordings are then processed depending on the research question. Most of the time, they are transcribed orthographically and phonemically or phonetically. Orthographic and phonemic transcriptions can be semi-automatised, implying different speech-to-text tools and dictionaries translating an orthographic transcription into phonemic symbols, mainly in the International Phonetic Alphabet (IPA). Transcriptions can also be done manually and should, in any case, be double-checked. In some cases, transcription does not need to be done before the next step when the recording is controlled and consists of specific sounds.

Acoustic analyses can be led in computer software such as Praat (Boersma & Weenink, 2022). This computer program allows for studying several acoustic parameters: pitch

(acoustic correlation of tone and intonation), intensity and loudness, duration, formant frequencies (*i.e.*, high energy frequency peaks in the spectrum, corresponding to resonances in the vocal tract), voice quality (jitter, shimmer, harmonic-to-noise ratio), cepstral information, and others (for more information, please consult <https://www.fon.hum.uva.nl/praat/>). It is also possible to automatise these measures for big corpora studies.

5.1.4 Voice quality measures

Acoustic parameters also include other factors that may require more advanced techniques. While vocal fold vibrations can be indirectly assessed in Praat (Boersma & Weenink, 2022), other instruments allow more direct observation. Electroglottograph (EGG) provides information on vocal folds abduction and adduction by measuring variations of the contact area between vocal folds during their vibration linked to voice production (Rothenberg, 1992). It is a non-invasive method where electrodes are placed around the neck (one or two electrodes on each side of the larynx) and fixed using a collar. A low-amperage current passes between the electrodes placed on each side of the thyroid cartilage. EGG is used to study vocal folds vibration patterns with no artefacts of vocal tract or aerodynamic noise influence. More precisely, it allows to obtain values for fundamental frequency (f_0), f_0 variation, jitter (f_0 instability measure reflecting the control of vocal folds vibration), shimmer (amplitude variation reflecting the glottal resistance), relative average perturbation (mean variation in consecutive periods), open quotient (OQ; percentage of each glottal cycle where the glottis is open), or closing quotient (CQ; ratio between the time closed and the complete glottal cycle allowing for studying glottal hypo- or hyperadduction).

5.1.5 Aerodynamics

Voice and voice quality studies can be completed using aerodynamic air pressure and airflow data. Indeed, voice can only be produced during exhalation when the air pressure is sufficient to overcome the abduction of the vocal folds. It hence plays an essential role in speech since appropriate air pressure is necessary for voice production and, in phonetics, it allows studying phenomena such as stop consonant production. The subglottal air pressure is measured in hectopascals (hPa) or centimetres of water (cm H₂O), and several techniques can be used for its estimation. The most direct procedures are invasive and consist of percutaneous, translaryngeal, or intraesophageal punctures of the cricothyroid membrane (Gross et al.,

2006; Plant & Hillel, 1998) or pressure sensors inserted into the trachea (Mehta et al., 2021). Other, less direct but also less invasive techniques include interpolation from intraoral air pressure measurements (Demolin et al., 1997), airway interruption method (the subject is tasked to produce specific sound sequences while holding a tubular device measuring airflow and subglottic pressure; Smitheran & Hixon, 1981), oesophageal balloon (Lieberman, 1968), respiratory inductance plethysmography bands (Salomoni et al., 2016), and others. Regarding airflow, its measures are mostly non-invasive and help assess neuromuscular disorders of the larynx and velum and voice disorders (Jiang et al., 2016). Airflow measurement techniques primarily include pneumotachographs (sensors measuring the gaseous exchange).

5.1.6 Articulation

In other cases, the study's object requires analysing articulatory behaviour rather than purely vocal parameters. Two main vocal tract areas can be assessed: laryngeal and supraglottic articulatory behaviours.

5.1.6.1 Laryngeal level

Studying articulation at the laryngeal level often implies focusing on vocal folds, adductions, and abductions. When direct observations are needed, they are done using the EGG, as mentioned earlier, or a laryngeal ultrasound, which allows for studying the laryngeal height (Moisik et al., 2014). Another tool for glottal activity observation, laryngoscopy, allows for capturing videos of vocal fold movement and is widely used to assess laryngeal behaviour in singers (Lam Tang et al., 2008), voice disorders (Bonilha et al., 2012), human beatbox (Dehais Underdown et al., 2019), and others. Unlike EGG, laryngoscopy is an invasive method primarily used for medical purposes and providing direct insight into vocal fold activity.

5.1.6.2 Supra-laryngeal level

At the supraglottic (or supra-laryngeal level), articulatory studies are interested in tongue, lips, velum, and jaw activity. As in aerodynamic studies, several tools can be used, depending on the research question. First, number of articulatory parameters can be obtained, though indirectly, in a computer program for acoustic analysis with fundamental frequency, formant frequencies, and other (spectral and cepstral) measures. When a direct observation is needed, it is possible to use ultrasounds (for tongue observation), velotrace (velum), video (lips, jaw), optotrace (visible articulators), electromagnetic articulography (lips, jaw, tongue), or

magnetic resonance imaging (MRI, the most complete observation). Analysed with appropriate software, these techniques permit investigations of the supraglottic articulatory speed, velocity, gesture duration, gesture anticipation, and others.

Table 2: Measurement techniques for speech production in experimental phonetics and phonology, and their invasiveness (Yes or No)

Research question		Method	Invasiveness
Acoustic analysis		Voice recording	N
Voice quality		Voice recording	N
		Electroglottography	N
Aerodynamics	Subglottal air pressure	Punctures	Y
		Airway interruption	N
		Intra-oral air pressure	N
		Oesophageal balloon	N
		Pletysmography	N
	Airflow	Pneumotachography	N
Laryngeal activity		Voice recording	N
		Electroglottography	N
		Ultrasounds	N
		Laryngoscopy	Y
Supraglottic activity		Voice recording	N
		Ultrasounds	N
		Electromagnetic articulography	Y
		MRI	Y

6. RECOMMENDATIONS FOR PRACTICE

Protocol design is the most challenging part in an experimental study. The difficulty increases when one manipulates speech samples, which involves recording, synthesising, or manipulating sound samples, and not only their written counterpart. In this section, we

provide guidance on critical steps in experimental research. As for the choice of the experimental tasks, which is a crucial issue, see the previous section. An experimental procedure combining production and perception tasks will offer a comprehensive view since ‘an operable function must, by definition, be contrastively encoded through production, and reliably decoded through perception’ (Xu, 2011, p. 90).

6.1 Testable hypotheses

Experimentation is based on a phonological or phonetic model to be tested. A model comes from a thorough observation and meaningful analysis of the phenomenon under scrutiny (Barbosa, 2012: 1), and experimental control ‘is about how to guarantee systematic identification and separation of the factors that may significantly contribute to what is under observation’ (Xu, 2011: 88). Hypotheses are linked to the functions of phonetic and phonological units. For illustration purposes, at the phonemic level, the example comes from Thai, a language where vowel length has a contrastive value (Abramson & Reo, 1990). An experimental procedure is needed to elucidate whether the underlying mechanism controls the relative duration of the articulations for the phonemes in contrast or whether short and long counterparts have somewhat different vowel qualities (timbre). The independent variable is the vowel duration (artificially lengthened or shortened without modifying vowel quality), and the dependent variable is the phoneme identification made by naïve listeners. At the prosodic level, a second example questions the relative effect of silent pauses and syntactic structure on the perception of boundaries in natural speech (Simon & Christodoulides, 2016). In the perceptual part of the experiment, naïve listeners are presented with samples of normal vs delexicalised speech and instructed to press a key whenever they perceive the end of “a group of words”. The controlled, independent variable is the syntactic structure, and the dependent variable is the boundary score assigned to each final-word syllable. Other acoustic variables are the pause duration or the final-word syllable relative length. Statistical analyses allow for studying the relative importance of acoustic and syntactic cues on the perception of a boundary. In any case, a clear delimitation of independent, dependent, and to-be-controlled variables is needed in the study design in a form that permits the application of statistical analysis (Xu, 2011).

6.2 Speakers' selection

Careful selection of speakers recorded for stimuli preparation is crucial to avoid bias in the experiment. Selection criteria depend on the study's objectives: native or non-native speakers, with standard or regional accents, naïve or professional, etc. It is difficult to control which aspect will be perceived as related to the speaker's gender, dialectal origin, or age; and to control which verbal or phonetic parameters will be decisive in the identification task (Simon et al., 2012).

6.3 Recording equipment

A digital recorder and a unidirectional microphone are usually used, and the choice of equipment can influence the acoustic analysis (Hansen & Farao, 2006). Making these recordings in a soundproof studio or in a quiet environment without surrounding noise is essential. The distance between the speaker and the microphone will influence certain acoustic parameters, such as intensity. With the pandemic, face-to-face recordings became impossible and remote recordings were used interactively or via self-recordings of subjects who were sent instructions beforehand.

Magistro (2021) tested four devices: a professional recorder and headset microphone, tablet, laptop, and Android phone. In conclusion, it is shown that Zencast software installed on a laptop is the most optimal software when professional equipment cannot be used, although the other devices can be used for the study of intonation.

6.3 Recruiting the participants

In a perception experiment, it is important to check several characteristics of the participants that could bias the task results. The participants must not have a hearing impairment. Those who suffer from hearing impairment should usually be excluded from the results (Thomas, 2002, p. 134). The level of expertise (lay speakers or experts) or the linguistic background (L1 or L2 speakers) should also be checked (Shen & Xatt, 2015). For some simple annotation tasks, recruiting non-native participants may be advantageous (Hasewaga-Johnson et al., 2015). In speech prosody studies, one often asks about the musical skills of the participant (Cason et al., 2015) since it may correlate with their ability to perceive fundamental frequency, accentuation or rhythm. Instructing participants to use quality listening equipment

(headphones) and pass the experiment in a quiet room without interrupting the trial is also important. As for the number of participants to be recruited, it depends on the statistical analyses to be carried out.

6.4 Presentation of stimuli

The number and duration of the stimuli should be controlled. They should be presented randomly to control for habituation or learning effects. Often, samples include a carrier sentence (i.e., a sentence where the stimulus of interest is embodied, such as ‘Say baba again. Say papa again. Say tata again.’). When an identification task (speaker’s origins, genre, speaking style) is required, it has been shown that the length of the samples has little influence on the results (Vieru et al., 2011) and that excerpts of 6 to 20 seconds are sufficient in most cases to allow listeners to make their judgement. A session should not last longer than 10 or 15 minutes. Otherwise, the task may be given up, especially if it is done remotely or without compensation.

6.5 Statistical analysis

Statistical tests to be used on the dataset should be determined simultaneously with the experimental protocol. It should be clear which variables will be studied, what will be considered dependent and independent variables, and what their nature is (qualitative, quantitative, ordinal). If interactions between variables are studied, this should be defined beforehand. The number of participants and the number of trials are essential since they also determine the use or not of a parametric test. Finally, it should be known if the same population will be asked to repeat a specific activity in a given lapse of time (paired samples, e.g. in longitudinal studies) or if one single participation is required (independent samples, typically group comparisons).

7. FUTURE PERSPECTIVES

In the experimental approach, phonetics and phonology are not to be conceived of as two separate disciplines (Ohala, 1990). Whether in production or perception, the analysis of observed acoustic or articulatory phenomena always depends on the consideration of their realisation within a domain (syllable, word, utterance, turn of speech...) and of their function in each linguistic system (Cangemi & Baumann, 2019).

Experimental phonetics and phonology have developed a wide range of scientific methods that allow testing theories about speech production and perception. Interdisciplinarity remains crucial since the theories to be tested derive from psychology of language, cognitive linguistics, articulatory phonology, sociolinguistics, discourse analysis, etc. Interdisciplinarity is also fruitful at the methodological level, in the combination of experimental linguistics and corpus linguistics (Gilquin & Gries, 2014), or experimental linguistics and computer modelling (Xu, 2011, p. 95).

Statistical analysis is an indispensable step in any experimental research. New methods of speech data analysis are emerging and are discussed: ‘Advances in computational power have made new analytical approaches possible, and the use of open access software such as R [...] increases the speed with which new statistical methods are shared both within our field and across disciplines. As accessibility to these methods increases, more and more people within linguistics employ increasingly complex analytical techniques (Roettger et al., 2019: 1). These developments require that the scripts used to perform the analyses be made available and open to critical discussion.

Finally, the movement towards open science forces to make primary data (sound or audio-visual recordings) and secondary data (acoustic measurements, transcripts, annotations, etc.) accessible, the issue being the cumulative and reproducible nature of the research results. ‘Data sharing is a crucial and necessary part of responsible conduct in research.’ (Garellek et al., 2020, p. 3). This time-consuming requirement is difficult to guarantee as the laws protecting personal data have recently been strengthened and academic institutions sometimes apply them in a very restrictive way. Nevertheless, what phonetic experimenters need to be prepared for is to make their data available to the scientific community and usable by other researchers. This activity of curating and archiving data, already in use for a long time in corpus phonology (Durand et al., 2014), is becoming an integral part of the experimental phonetics and phonology publication activity itself (as evidenced by the new editorial policy of *the Journal of Phonetics*, Cho, 2021).

8. FURTHER READINGS

Cohn, A. C., Fougeron, C., Huffman, M.K. (2011). *The Oxford Handbook of Laboratory Phonology*. Oxford University Press.

This manual details a large number of concepts and methods useful for the study of language production, perception and use, at the segmental and suprasegmental levels.

Durand, J., Gut, U., Kristoffersen, G. (2014). *The Oxford Handbook of Corpus Phonology*. Oxford: Oxford University Press.

Another important book on phonology, presenting different types of corpora used in experimental phonology.

Katz, W.F., Assmann P.F. (2019) *The Routledge Handbook of Phonetics*. Routledge.

The Handbook includes information on physiological aspects of speech, on theories and models used in phonetics, as well as information on real-life applications of phonetic sciences.

Knight R.A., Setter J. (2022) *The Cambridge Handbook of Phonetics*. Cambridge University Press.

This book provides important insights on sound and prosody production, audition, and perception.

9. RELATED TOPICS

acoustics; articulation; pathological speech; prosody; speech; voice.

10. REFERENCES

- Abramson, A.S. & Reo, N. (1990). Distinctive vowel length: duration vs. spectrum in Thai, *Journal of Phonetics*, 18(2), 79–92. doi:[10.1016/S0095-4470\(19\)30395-X](https://doi.org/10.1016/S0095-4470(19)30395-X).
- Alday, P.M. (2019). M/EEG analysis of naturalistic stories: a review from speech to language processing, *Language, Cognition and Neuroscience*, 34(4), 457–473. doi:[10.1080/23273798.2018.1546882](https://doi.org/10.1080/23273798.2018.1546882).
- Anansiripinyo, T. & Onsuwan, C. (2019). Acoustic-phonetic characteristics of Thai filled pauses in monologues. In *The 9th Workshop on Disfluency in Spontaneous Speech (DiSS 2019)*. Budapest: Budapest, 51–54. doi:[10.21862/diss-09-014-anan-onsu](https://doi.org/10.21862/diss-09-014-anan-onsu).
- Avelino, H., Coler, M. & Wetzels, L. (Eds.). (2016). *The Phonetics and Phonology of Laryngeal Features in Native American Languages*. Leiden: Brill. Available at: <https://brill.com/view/title/32268> (Accessed: 14 June 2022).
- Awan, S. N. (2008). Instrumental Analysis of Phonation. In M. J. Ball, M. R. Perkins, N. Miller, & S. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 344–359). Blackwell Publishing Ltd. <https://doi.org/10.1002/9781444301007.ch21>

- Bänziger, T. & Scherer, K.R. (2005). The role of intonation in emotional expressions, *Speech Communication*, 46(3–4), 252–267. doi:[10.1016/j.specom.2005.02.016](https://doi.org/10.1016/j.specom.2005.02.016).
- Boë, L.-J. & Vilain, C. (Eds.). (2011). *Un siècle de phonétique expérimentale : Fondation et éléments de développement, Hommage à Théodore Rosset et John Ohala*. ENS Editions. Lyon (Langages). Available at: <http://catalogue-editions.ens-lyon.fr/fr/livre/?GCOI=29021100097800> (Accessed: 11 October 2021).
- Boersma, P. & Weenink, D. (2022). *Praat: doing phonetics by computer (Version 6.2.12)*. Available at: <http://www.praat.org/>.
- Bonilha, H. S., White, L., Kuckhahn, K., Gerlach, T. T., & Deliyski, D. D. (2012). Vocal fold mucus aggregation in persons with voice disorders. *Journal of Communication Disorders*, 45(4), 304–311. <https://doi.org/10.1016/j.jcomdis.2012.03.001>
- Boula de Mareüil, Ph., & Vieru-Dimulescu. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, 63, 247–267.
- Brentari, D. (2019). *Sign Language Phonology* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781316286401>
- Brognaux, S. & Drugman, T. (2014). Phonetic variations: Impact of the communicative situation, in *Proc. 7th International Conference on Speech Prosody 2014*. *Speech Prosody* 7, 428–432. doi:[DOI: 10.21437/SpeechProsody.2014-72](https://doi.org/10.21437/SpeechProsody.2014-72).
- Caccia, M. & Lorusso, M.L. (2019). When prosody meets syntax: The processing of the syntax-prosody interface in children with developmental dyslexia and developmental language disorder, *Lingua*, 224, 16–33. doi:[10.1016/j.lingua.2019.03.008](https://doi.org/10.1016/j.lingua.2019.03.008).
- Campbell, L. & Grondona, V. (2012). *The Indigenous Languages of South America: A Comprehensive Guide*, *The Indigenous Languages of South America*. Berlin, Boston: De Gruyter Mouton. doi:[10.1515/9783110258035](https://doi.org/10.1515/9783110258035).
- Cangemi, F. & Baumann, S. (2020). Integrating phonetics and phonology in the study of linguistic prominence, *Journal of Phonetics*, 81, 100993. doi:[10.1016/j.wocn.2020.100993](https://doi.org/10.1016/j.wocn.2020.100993).
- Cason, N., Astésano, C. & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception, *Acta Psychologica*, 155, 43–50. doi:[10.1016/j.actpsy.2014.12.002](https://doi.org/10.1016/j.actpsy.2014.12.002).

- Cebrian, J. (2021). Perception of English and Catalan vowels by English and Catalan listeners: A study of reciprocal cross-linguistic similarity, *The Journal of the Acoustical Society of America*, 149(4), 2671–2685. doi:[10.1121/10.0004257](https://doi.org/10.1121/10.0004257).
- Cho, T. (2021). Where we are at: Impact, special collections, open science and registered report at the Journal of Phonetics, *Journal of Phonetics*, 89. doi:<https://doi.org/10.1016/j.wocn.2021.101113>.
- Cole, J., Mahrt, T. & Roy, J. (2017). Crowd-sourcing prosodic annotation, *Computer Speech & Language*, 45, 300–325. doi:[10.1016/j.csl.2017.02.008](https://doi.org/10.1016/j.csl.2017.02.008).
- Dehais Underdown, A., Buchman, L. & Demolin, D. (2019). Acoustico-Physiological coordination in the Human Beatbox: A pilot study on the beatboxed Classic Kick Drum, in *19th International Congress of Phonetic Sciences*. Melbourne, Australia. Available at: <https://hal.archives-ouvertes.fr/hal-02284132> (Accessed: 14 June 2022).
- Demolin, D., Giovanni, A., Hassid, S., Heim, C., Lecuit, V., & Soquet, A. (1997). Direct and Indirect Measurements of Subglottic Pressure, in *LARYNX 1997*. Marseille, 69–72. Available at: https://www.isca-speech.org/archive_open/larynx_97/lar7_069.html (Accessed: 14 June 2022).
- Demolin, D. (2012). Experimental methods in phonology, *TIPA. Travaux interdisciplinaires sur la parole et le langage* [Preprint], (28). Available at: <http://tipa.revues.org/162> (Accessed: 9 January 2013).
- Demolin, D. & Chabiron, C. (2013). Clicks, stop bursts, vocoids and the timing of articulatory gestures in Rwanda, in *Phonetics and Phonology of Sub-Saharan Languages*. Johannesburg, South Africa. Available at: <https://hal.archives-ouvertes.fr/hal-00834241> (Accessed: 14 June 2022).
- Didirková, I., Crible, L. & Simon, A.C. (2019). Impact of Prosody on the Perception and Interpretation of Discourse Relations: Studies on “Et” and “Alors” in Spoken French, *Discourse Processes*, 56(8), 619–642. doi:[10.1080/0163853X.2018.1528963](https://doi.org/10.1080/0163853X.2018.1528963).
- Didirková, I., Le Maguer, S. & Hirsch, F. (2021). An articulatory study of differences and similarities between stuttered disfluencies and non-pathological disfluencies, *Clinical Linguistics & Phonetics*, 35(3), 201–221. doi:[10.1080/02699206.2020.1752803](https://doi.org/10.1080/02699206.2020.1752803).
- Dogil, G., Ackermann, H., Grodd, W., Haider, H., Kamp, H., Mayer, J., Riecker, A., & Wildgruber, D. (2002). The speaking brain: A tutorial introduction to fMRI experiments in the production of speech, prosody and syntax. *Journal of Neurolinguistics*, 15(1), 59–90. [https://doi.org/10.1016/S0911-6044\(00\)00021-X](https://doi.org/10.1016/S0911-6044(00)00021-X)

- Dufour, S., Mirault, J. & Grainger, J. (2022). Investigating the locus of transposed-phoneme effects using cross-modal priming, *Acta Psychologica*, 226, 103578. doi:[10.1016/j.actpsy.2022.103578](https://doi.org/10.1016/j.actpsy.2022.103578).
- Durand, J., Laks, B. & Lyche, C. (2014). French Phonology from a Corpus Perspective: perspective: The PFC Programme. In Durand, J., Gut, U. & Kristoffersen, G. (Eds.). *The Oxford Handbook of Corpus Phonology* (pp. 86–497). Oxford: Oxford University Press, doi:[10.1093/oxfordhb/9780199571932.013.015](https://doi.org/10.1093/oxfordhb/9780199571932.013.015).
- Fischer, K. (Ed.). (2021). *Approaches to Discourse Particles*. Leiden: Brill. Available at: <https://brill.com/view/title/23248> (Accessed: 14 June 2022).
- Hansen, G. F. & Pharaoh, N. (2006). Microphones and measurements. In Ambrazaitis, G. & Schötz, S. (Eds.). *Working Papers. Proceedings from Fonetik 2006* (pp. 49-52). Lund: Lund University, Centre for Languages and Literature. Available at: https://www.academia.edu/864240/Microphones_and_measurements (Accessed: 14 December 2021).
- Garellek, M., Gordon, M., Kirby, J., Lee, W.-S., Michaud, A., Mooshammer, C., Niebuhr, O., Recasens, D., Roettger, T. B., Simpson, A., & Yu, K. M. (2020). Toward open data policies in phonetics: What we can gain and how we can avoid pitfalls. *Journal of Speech Science*, 9(1), 3.
- Georgiades, M. & Sfakianaki, A. (2004). Syntax interacts with prosody: the VOS order in Greek, *Lingua*, 114(7), 935–961. doi:[10.1016/S0024-3841\(03\)00099-8](https://doi.org/10.1016/S0024-3841(03)00099-8).
- Gilquin, G. & Gries, S.T. (2009). Corpora and experimental methods: A state-of-the-art review, *Corpus Linguistics and Linguistic Theory*, 5(1), 1–26.
- Goldman, J. P., Pršir, T., Christodoulides, G., Simon, A. C., & Auchlin, A. (2014). Phonogenre Identification: A Perceptual Experiment with 8 Delexicalised Speaking Styles, *Nouveaux Cahiers de Linguistique Française*, 31, 51–62.
- Gonen, E., Livnat, Z. & Amir, N. (2015). The discourse marker axshav (“now”) in spontaneous spoken Hebrew: Discursive and prosodic features, *Journal of Pragmatics*, 89, 69–84. doi:[10.1016/j.pragma.2015.09.005](https://doi.org/10.1016/j.pragma.2015.09.005).
- Gras, P. & Elvira-García, W. (2021). The role of intonation in Construction Grammar: On prosodic constructions, *Journal of Pragmatics*, 180, 232–247. doi:[10.1016/j.pragma.2021.05.010](https://doi.org/10.1016/j.pragma.2021.05.010).

- Grigos, M.I. (2009). Changes in articulator movement variability during phonemic development: a longitudinal study, *Journal of speech, language, and hearing research: JSLHR*, 52(1), 164–177. doi:[10.1044/1092-4388\(2008/07-0220\)](https://doi.org/10.1044/1092-4388(2008/07-0220)).
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm, *Perception & Psychophysics*, 28(4), 267–283. doi:[10.3758/bf03204386](https://doi.org/10.3758/bf03204386).
- Gross, R. D., Steinhauer, K. M., Zajac, D. J., & Weissler, M. C. (2006). Direct Measurement of Subglottic Air Pressure While Swallowing: *The Laryngoscope*, 116(5), 753–761. <https://doi.org/10.1097/01.mlg.0000205168.39446.12>
- Guenther, F.H. (2016) *Neural Control of Speech*. Cambridge, MA, USA: MIT Press.
- Hasegawa-Johnson, M., Cole, J., Jyothi, P., & Varshney, L. R. (2015). Models of dataset size, question design, and cross-language speech perception for speech crowdsourcing applications. *Laboratory Phonology*, 6(3–4), 381–431. <https://doi.org/10.1515/lp-2015-0012>
- Hayward, K. (2000). *Experimental Phonetics: An Introduction*. Routledge.
- Hsiao, Y.E. (2020). The syntax-prosody competition: Evidence from adjunct prosodic parsing in iGeneration Taiwanese, *Lingua*, 237, 102805. doi:[10.1016/j.lingua.2020.102805](https://doi.org/10.1016/j.lingua.2020.102805).
- Huang, Y. T., Newman, R. S., Catalano, A., & Goupell, M. J. (2017). Using prosody to infer discourse prominence in cochlear-implant users and normal-hearing listeners. *Cognition*, 166, 184–200. <https://doi.org/10.1016/j.cognition.2017.05.029>
- Jacewicz, E., Fox, R.A. & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English, *The Journal of the Acoustical Society of America*, 128(2), 839–850. doi:[10.1121/1.3459842](https://doi.org/10.1121/1.3459842).
- Jiang, J. J., Hanna, R. B., Willey, M. V., & Rieves, A. (2016). The Measurement of Airflow Using Singing Helmet That Allows Free Movement of the Jaw. *Journal of Voice*, 30(6), 641–648. <https://doi.org/10.1016/j.jvoice.2015.07.018>
- Kim, S.-E. & Tilsen, S. (2022). An investigation of functional relations between speech rate and phonetic variables, *Journal of Phonetics*, 93, 101152. doi:[10.1016/j.wocn.2022.101152](https://doi.org/10.1016/j.wocn.2022.101152).
- King, H. & Ferragne, E. (2020). Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English, *Journal of Phonetics*, 80, 100978. doi:[10.1016/j.wocn.2020.100978](https://doi.org/10.1016/j.wocn.2020.100978).
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: U. of Pennsylvania Press.

- Ladefoged, P. (2004). Phonetics and phonology in the last 50 years, in *From Sound to Sense*. MIT.
- Lam Tang, J.A., Boliek, C.A. & Rieger, J.M. (2008). Laryngeal and respiratory behavior during pitch change in professional singers, *Journal of Voice: Official Journal of the Voice Foundation*, 22(6), 622–633. doi:[10.1016/j.jvoice.2007.04.002](https://doi.org/10.1016/j.jvoice.2007.04.002).
- Lawson, E. & Stuart-Smith, J. (2021). Lenition and fortition of /r/ in utterance-final position, an ultrasound tongue imaging study of lingual gesture timing in spontaneous speech, *Journal of Phonetics*, 86, p. 101053. doi:[10.1016/j.wocn.2021.101053](https://doi.org/10.1016/j.wocn.2021.101053).
- Lehner, K. & Ziegler, W. (2022). Indicators of Communication Limitation in Dysarthria and Their Relation to Auditory-Perceptual Speech Symptoms: Construct Validity of the KommPaS Web App, *Journal of speech, language, and hearing research*, 65(1), 22–42. doi:[10.1044/2021_JSLHR-21-00215](https://doi.org/10.1044/2021_JSLHR-21-00215).
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.
- Lieberman, P. (1968). Direct comparison of subglottal and esophageal pressure during speech, *The Journal of the Acoustical Society of America*, 43(5), 1157–1164. doi:[10.1121/1.1910950](https://doi.org/10.1121/1.1910950).
- van Lieshout, P., Merrick, G. & Goldstein, L. (2008). An Articulatory Phonology Perspective on Rhotic Articulation Problems: A Descriptive Case Study, *Asia Pacific Journal of Speech, Language and Hearing*, 11(4), 283–303. doi:[10.1179/136132808805335572](https://doi.org/10.1179/136132808805335572).
- Ma, B. & Zhuang, H. (2022). Prosody and sentencehood in Chinese: A case study with reference to Chinese-English comparison, *Lingua*, 270, 103217. doi:[10.1016/j.lingua.2021.103217](https://doi.org/10.1016/j.lingua.2021.103217).
- Magistro, G. (2021). Speech prosody and remote experiments: a technical report, *arXiv:2106.10915 [cs, eess]* [Preprint]. Available at: <http://arxiv.org/abs/2106.10915> (Accessed: 30 September 2021).
- Mehta, D. D., Kobler, J. B., Zeitels, S. M., Zañartu, M., Ibarra, E. J., Alzamendi, G. A., Manriquez, R., Erath, B. D., Peterson, S. D., Petrillo, R. H., & Hillman, R. E. (2021). Direct Measurement and Modeling of Intraglottal, Subglottal, and Vocal Fold Collision Pressures during Phonation in an Individual with a Hemilaryngectomy. *Applied Sciences*, 11(16), 7256. <https://doi.org/10.3390/app11167256>

- Merlo, S. & Mansur, L.L. (2004). Descriptive discourse: topic familiarity and disfluencies, *Journal of Communication Disorders*, 37(6), 489–503.
doi:[10.1016/j.jcomdis.2004.03.002](https://doi.org/10.1016/j.jcomdis.2004.03.002).
- Michaud, A. (2012). The Complex Tones of East/Southeast Asian Languages: Current Challenges for Typology and Modelling.
- Michelas, A. & Dimperio, M. (2012). When syntax meets prosody: Tonal and duration variability in French Accentual Phrases, *Journal of Phonetics*, 40(6), 816–829.
doi:[10.1016/j.wocn.2012.08.004](https://doi.org/10.1016/j.wocn.2012.08.004).
- Moisik, S.R., Lin, H. & Esling, J.H. (2014). A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS), *Journal of the International Phonetic Association*, 44(1), 21–58.
doi:[10.1017/S0025100313000327](https://doi.org/10.1017/S0025100313000327).
- Moniz, H., Batista, F., Mata, A. I., & Trancoso, I. (2014). Speaking style effects in the production of disfluencies. *Speech Communication*, 65, 20–35.
<https://doi.org/10.1016/j.specom.2014.05.004>
- Moradi, S., Lidestam, B., Saremi, A., & Rönnberg, J. (2014). Gated auditory speech perception: Effects of listening conditions and cognitive capacity. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00531>
- Nicolaidis, K. & Baltazani, M. (2014). The Greek Rhotic in /rC/ Sequences: An Acoustic and Electropalatographic Study. In Lavidas, N., Alexiou, T., & Sougari, A.M. (Eds.). *Major Trends in Theoretical and Applied Linguistics 1. Selected Papers from the 20th ISTAL* (pp. 157-176). De Gruyter Open Poland. doi:[10.2478/9788376560762.p22](https://doi.org/10.2478/9788376560762.p22).
- Ohala, J.J. (1987). Experimental Phonology in *Proceedings of the Annual Meeting, Berkeley Linguistic Society*. Berkeley, California, 207–222.
- Ohala, J.J. (1990). There is no interface between phonology and phonetics: a personal view, *Journal of Phonetics*, 18(2), 153–171. doi:[10.1016/S0095-4470\(19\)30399-7](https://doi.org/10.1016/S0095-4470(19)30399-7).
- Ohala, J.J. (2010) The Relation between Phonetics and Phonology, in Hardcastle, W., Laver, J. & Gibbon, F.E. (Eds.). *The Handbook of Phonetic Sciences* (2nd ed.) (pp. 653-677). John Wiley & Sons, Ltd.
- Ohala, J.J. & Jaeger, J.J. (1986). *Experimental Phonology*. Oxford: Academic Press.
- Overby, M. & Caspari, S.S. (2015). Volubility, consonant, and syllable characteristics in infants and toddlers later diagnosed with childhood apraxia of speech: A pilot study, *Journal of Communication Disorders*, 55, 44–62. doi:[10.1016/j.jcomdis.2015.04.001](https://doi.org/10.1016/j.jcomdis.2015.04.001).

- Perlman, M. & Benitez, N.J. (2008). Talking Fast: The Use of Speech Rate as Iconic Gesture, in. *9th Conference on Conceptual Structure, Discourse, & Language (CSDL9)*, Rochester, NY: Social Science Research Network. Available at: <https://ssrn.com/abstract=1299284> (Accessed: 9 September 2013).
- Plant, R.L. & Hillel, A.D. (1998). Direct measurement of subglottic pressure and laryngeal resistance in normal subjects and in spasmodic dysphonia, *Journal of Voice: Official Journal of the Voice Foundation*, 12(3), 300–314. doi:[10.1016/s0892-1997\(98\)80020-9](https://doi.org/10.1016/s0892-1997(98)80020-9).
- Poupier, M. & Beňuš, Š. (2009). Phonetic properties of syllabic consonants in Slovak., *The Journal of the Acoustical Society of America*, 125(4), 2569–2569. doi:[10.1121/1.4783740](https://doi.org/10.1121/1.4783740).
- Prieto, P., Pugliesi, C., Borràs-Comes, J., Arroyo, E., & Blat, J. (2015). Exploring the contribution of prosody and gesture to the perception of focus using an animated agent. *Journal of Phonetics*, 49, 41–54. <https://doi.org/10.1016/j.wocn.2014.10.005>
- Przewozny, A., Viollain, C. & Navarro, S. (Eds.). (2020). *The Corpus Phonology of English: Multifocal Analyses of Variation*. Edinburgh: Edinburgh University Press. Available at: <https://www.jstor.org/stable/10.3366/j.ctv177th25> (Accessed: 15 June 2022).
- Rathcke, T., Lin, C.-Y., Falk, S., & Bella, S. D. (2021). Tapping into linguistic rhythm. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 12(1), 11. <https://doi.org/10.5334/labphon.248>
- Redcay, E., Haist, F. & Courchesne, E. (2008). Functional neuroimaging of speech perception during a pivotal period in language acquisition, *Developmental Science*, 11(2), 237–252. doi:[10.1111/j.1467-7687.2008.00674.x](https://doi.org/10.1111/j.1467-7687.2008.00674.x).
- Roettger, T.B., Winter, B. & Baayen, H. (2019). Emergent data analysis in phonetic sciences: Towards pluralism and reproducibility, *Journal of Phonetics*, 73, 1–7. doi:[10.1016/j.wocn.2018.12.001](https://doi.org/10.1016/j.wocn.2018.12.001).
- Rothenberg, M. (1992). A multichannel electroglottograph, *Journal of Voice*, 6(1), 36–43. doi:[10.1016/S0892-1997\(05\)80007-4](https://doi.org/10.1016/S0892-1997(05)80007-4).
- Rousselot, P.-J. (1924). *Principes de phonétique expérimentale*. Paris: Didier.
- de Ruiter, J.P., Bangerter, A. & Dings, P. (2012). The Interplay Between Gesture and Speech in the Production of Referring Expressions: Investigating the Tradeoff Hypothesis, *Topics in Cognitive Science*, 4(2), 232–248. doi:[10.1111/j.1756-8765.2012.01183.x](https://doi.org/10.1111/j.1756-8765.2012.01183.x).

- Salomoni, S., van den Hoorn, W. & Hodges, P. (2016). Breathing and Singing: Objective Characterization of Breathing Patterns in Classical Singers, *PLoS ONE*, 11(5), e0155084. doi:[10.1371/journal.pone.0155084](https://doi.org/10.1371/journal.pone.0155084).
- Scripture, E.W. (1902). *The elements of experimental phonetics*. Cambridge: Cambridge University Press.
- Sedhu, A. (2015) A Comparative Study of the Rhythm of Fulfulde and Hausa, *Journal of the Linguistic Association of Nigeria*, 20, 200–215.
- Shattuck-Hufnagel, S. (2019). Toward an (even) more comprehensive model of speech production planning, *Language, Cognition and Neuroscience*, 34(9), 1202–1213. doi:[10.1080/23273798.2019.1650944](https://doi.org/10.1080/23273798.2019.1650944).
- Shen, C. & Watt, D. (2015). Accent Categorisation by Lay Listeners: Which Type of “Native Ear” Works Better?, *York Papers In Linguistics (YPL2)*, 14. Available at: https://www.academia.edu/15470009/Accent_Categorisation_by_Lay_Listeners_Which_Type_of_Native_Ear_Works_Better (Accessed: 23 May 2022).
- Signorello, R., Hassid, S. & Demolin, D. (2017). Aerodynamic Features of French Fricatives, in *Proc. of Interspeech 2017. Interspeech 2017*, Stockholm, Sweden, 2267–2271. Available at: https://www.academia.edu/74129440/Aerodynamic_Features_of_French_Fricatives (Accessed: 15 June 2022).
- Simon, A. C., Hambye, P., Bardiaux, A., & Boula de Mareüil, P. (2012). Caractéristiques des accents régionaux en français : que nous apprennent les approches perceptives ?. In Simon, A.C. (Ed.). *La variation prosodique régionale en français* (pp. 27-40). Louvain-la-Neuve: De Boeck Supérieur. Available at: <https://www.cairn.info/la-variation-prosodique-regionale-en-francais--9782801116951-page-27.htm> (Accessed: 27 March 2020).
- Simon, A.-C. & Auchlin, A. (2001). Multi-modal, multi-focal : les “hors-phase” de la prosodie, in Cavé, C., Guaïtella, I., & Santi, S. (Eds.). *Oralité et gestualité. Interaction et comportements multimodaux dans la communication* (pp. 629-633). Paris: LHarmattan.
- Simon, A.C. & Christodoulides, G. (2016). Frontières prosodiques perçues : corrélats acoustiques et indices syntaxiques, *Langue française*, 191(3), 83–106.

- Smitheran, J.R. & Hixon, T.J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production, *The Journal of Speech and Hearing Disorders*, 46(2), 138–146. doi:[10.1044/jshd.4602.138](https://doi.org/10.1044/jshd.4602.138).
- Steinhauer, K., Abada, S. H., Pauker, E., Itzhak, I., & Baum, S. R. (2010). Prosody–syntax interactions in aging: Event-related potentials reveal dissociations between on-line and off-line measures. *Neuroscience Letters*, 472(2), 133–138. <https://doi.org/10.1016/j.neulet.2010.01.072>
- Stouten, F., Duchateau, J., Martens, J. P., & Wambacq, P. (2006). Coping with disfluencies in spontaneous speech recognition: Acoustic detection and linguistic context manipulation, *Speech Communication*, 48(11), 1590–1606. doi:[10.1016/j.specom.2006.04.004](https://doi.org/10.1016/j.specom.2006.04.004).
- Thomas, E.R. (2002). Sociophonetic Applications of Speech Perception Experiments, *American Speech*, 77(2), 115–147.
- Tomaschek, F., Arnold, D., Sering, K., Tucker, B. V., van Rij, J., & Ramscar, M. (2021). Articulatory Variability is Reduced by Repetition and Predictability. *Language and Speech*, 64(3), 654–680. <https://doi.org/10.1177/0023830920948552>
- Tyler, J. (2013). Prosodic correlates of discourse boundaries and hierarchy in discourse production, *Lingua*, 133, 101–126. doi:[10.1016/j.lingua.2013.04.005](https://doi.org/10.1016/j.lingua.2013.04.005)
- Vainio, M., Suni, A. S., Raitio, T., Nurminen, J., Järvikivi, J., & Alku, P. (2009). *New Method for Delexicalization and its Application to Prosodic Tagging for Text-to-Speech Synthesis*. <https://doi.org/10.21437/Interspeech.2009-514>
- Van Praet, W. & OGrady, G. (2018). The prosody of specification: Discourse intonational cues to setting up a variable, *Journal of Pragmatics*, 135, 87–100. doi:[10.1016/j.pragma.2018.07.013](https://doi.org/10.1016/j.pragma.2018.07.013).
- Vieru, B., de Mareüil, P.B. & Adda-Decker, M. (2011). Characterisation and identification of non-native French accents, *Speech Communication*, 53(3), 292–310. doi:[10.1016/j.specom.2010.10.002](https://doi.org/10.1016/j.specom.2010.10.002).
- Wagner, P., Trouvain, J. & Zimmerer, F. (2015). In defense of stylistic diversity in speech research, *Journal of Phonetics*, 48, 1–12. doi:[10.1016/j.wocn.2014.11.001](https://doi.org/10.1016/j.wocn.2014.11.001).
- Wang, J., Zhu, Y., Chen, Y., Mamat, A., Yu, M., Zhang, J., & Dang, J. (2020). An Eye-Tracking Study on Audiovisual Speech Perception Strategies Adopted by Normal-Hearing and Deaf Adults Under Different Language Familiarities. *Journal of Speech*,

Language, and Hearing Research, 63(7), 2245–2254.

https://doi.org/10.1044/2020_JSLHR-19-00223

- Warner, H. J., Whalen, D. H., Harel, D., & Jackson, E. S. (2022). The effect of gap duration on the perception of fluent versus disfluent speech. *Journal of Fluency Disorders*, 71, 105896. <https://doi.org/10.1016/j.jfludis.2022.105896>
- Wells, B., & Whiteside, S. (2008). Prosodic Impairments. In M. J. Ball, M. R. Perkins, N. Miller, & S. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 549–567). New-Jersey : Blackwell Publishing Ltd. <https://doi.org/10.1002/9781444301007.ch34>
- Wu, Y., Didirkova, I. & Simon, A.-C. (2022). Disfluences en parole continue en français : paramètres prosodiques des pauses pleines et des allongements vocaliques, in *Actes des 34e Journées d'Études sur la Parole. 34e Journées d'Études sur la Parole - JEP2022*. Available at: <https://dial.uclouvain.be/pr/boreal/object/boreal:260873> (Accessed: 15 June 2022).
- Xu, Y. (2011). Speech Prosody: A Methodological Review, *Journal of Speech Sciences*, 1(1), 85–115.
- Zerbian, S., & Barnard, E. (2008). Phonetics of intonation in South African Bantu languages, *Southern African Linguistics and Applied Language Studies*, 26(2), 235-254.