HONOURING HISTORICAL FACTS: THE CASE OF INTONATIONAL DOWNTRENDS

Jan Volín, Pavel Šturm
Institute of Phonetics, Faculty of Arts, Charles University, Czech Republic
jan.volin@ff.cuni.cz

Abstract: Various researchers of the past are honoured in citations as the very first who described or discovered certain scientifically important phenomena. However, the illusion of “our times being superior to past” causes certain myopia: we often do not go back enough to identify and appreciate the true pioneers. Generally, we tend to underestimate what our ancestors already knew and did, and we overestimate the contribution of our own era. The present paper demonstrates this predicament with the case of intonational downtrends, i.e., gradual lowering of phonologically equivalent melodic targets within an utterance. The first report of the phenomenon is often attributed to K. L. Pike, sometimes even to later intonologists. We, on the other hand, document much older accounts of intonational downtrends, and offer some tentative explanations as to why the earlier research was ignored.

1 Introduction

1.1 Rationale

Science as a way of human learning about the world possesses several crucial attributes. One of them is cumulativeness. Researchers who expand human knowledge always build on what has been achieved before them.

Many lay people believe that in terms of scientific endeavour our ancestors were somehow underdeveloped or even backward. This cognitive illusion stems from several sources. First, it is the state of technology that leads to a childish contention like “our ancestors did not have our ‘wonderful gadgets’, therefore, they must have been intellectually deficient”. This way of reasoning not only confuses innovations with science, but also ignores the fact that innovations largely benefit from science, and share the cumulativeness feature with it. Second, when teaching children to be grateful, there is a general tendency to portray the current state of affairs as advanced in comparison with backslidden or even degenerated past. Third, it is often pointed out how the outdated scientific views of the past were mistaken, but we do not realize that our current views will also be judged as outdated and mistaken one day. Moreover, what seems clearly wrong from today’s perspective in the science of the past actually constitutes the essential precondition of our current knowledge, so even the errors should be considered with gratitude and respect. It follows that one does not have to be a trained philosopher to appreciate the past as the essential prerequisite of the present.

Current conventions of empiricist discourse require accurate and unambiguous referencing whenever non-trivial facts, propositions, assertions, or quantitative results are presented in research reporting (see, e.g., American Psychological Association [1]). This seemingly simple requirement causes a surprisingly large number of problems. Our present paper demonstrates one of them: attributing certain suggestions, findings or primacy to the wrong person. The reasons for this can be many – anything from good faith in a source, through plain ignorance up to ill intentions.

It has to be stressed, however, that our chief objective in the present study is not to criticize blunders of our colleagues, because we realize only too well that no one is really immune to
slip-ups. Instead, we wish to expound some interesting facts about a feature of speech melody description that has been known for a longer time than various authors usually report.

1.2 Intonation and downtrends

Intonation itself was largely outside the focus of phonetic research until about 1960s. Although there were many inspired works from long before that (e.g., [10], [25], [2]), speech prosody was overshadowed by segmental issues. This can be easily documented by the lists of topics at the first three International Congresses of Phonetic Sciences (1932 in Amsterdam, 1935 in London, 1938 in Ghent), where papers related to melody of speech occupied an extremely small part of the space. Rossi [28: 2] speaks about an “exhaustive bibliography covering the years from 1900 to 1973” in which Di Cristo collected almost 4400 studies on speech prosody. When Rossi ordered them by their dates and split in the middle, he found out that the first half of them needed over 60 years to be written, while the other half was generated within just over 6 years. The turning point was 1967 (the year of the 6th ICPhS in Prague; see also [30] for similar comparisons between 1967 and 2011).

Various hypotheses could be stipulated concerning the unbalanced division of attention in the study of sound patterns in speech. First, it is true across all scientific disciplines that people initially deal with the obvious, and only later arrive at phenomena that are less accessible to conscious observation. Thanks to alphabetic spelling systems dominating the learned space in Europe and America, the segmental speech sounds, i.e., consonants and vowels, have always been the most salient constituents of the speech signal in terms of conscious accessibility. When, for instance, Kingdon [19: xxv] explains the difference between kinetic and static tones, he likens them to diphthongs and monophthongs as if he presumed that those would be better known to his readers.

Second, the linguistic theory of the early 20th century was deeply interconnected with semiotics – the general theory of signs. Since various prosodic units and features did not fit the dogma of a discrete and arbitrary sign, researchers were less keen on studying these. Even today, however, no agreement has been reached on the essential nature of the linguistic sign.

Third, the heavy burden of exclusive anthropocentrism in the early stages of linguistic endeavour prevented scholars from seeing language in its entirety. Rather than realizing that language is a complex system that enables communication of a wide range of meanings, conservative linguists limited their object of inquiry to lexicon and grammar. Various fundamental parts of the systems were labelled ‘paralinguistic’ and left to be studied by ‘someone else’. Yet, there were voices already in the 1930s that the representational, affective and applicative components of the meaning form a unity that must be taken seriously [5]. (It is natural that a linguist must specialize in, let’s say, morphonology, but to say that only this auxiliary discipline is the true linguistic undertaking would be deeply damaging to the whole discipline.)

It is becoming clearer and clearer nowadays that various features of intonation generate meanings of utterances the same way lexical items do, i.e., through interaction with contexts. Just like lexical items, intonational features possess potentials of meanings, and only through interacting with the concrete layers of the context the real meaning or sense is produced. Thus, intonation is not ‘added’ to words. It is selected together with the words, and it is uttered to interact with the context and to generate the intended meaning. No features of intonation should be a priori excluded from the language structure just because the current state of knowledge is unable to see their role in communication [22: 323].

As stated above, these ideas began entering the general approaches to speech communication extensively in the mid-1960s. At about the same time larger attention was becoming oriented towards a specific feature of $F_0$ contours: a gradual downtrend of $F_0$ values throughout utterances. This trend can be expressed as $\Delta F_0/\Delta t < 0$, and illustrated with linear regression lines throughout $F_0$ contours (e.g., Figure 1). More often, however, it is emphasized that the trend
should be studied as the gradual lowering of the *equivalent melodic events* like peaks or valleys. Rather than linear regression fit, researchers may use toplines (connecting peaks), baselines (connecting valleys), or even simple curves constructed with prescribed procedures.

The first significant public presentation of the phenomenon to larger audiences seems to have happened at the 5th International Congress of Acoustics in 1965 in Belgium (as alleged by ’t Hart et al. [33: 123]). A group of Dutch phoneticians affiliated with the Institute for Perception Research (IPO) in Eindhoven suggested a name for the phenomenon: they called it *declination*. They also demonstrated its perceptual relevance in Dutch utterances. Further research has added similar concepts under the labels such as downstep, downdrift, final lowering, etc. The umbrella term *downtrends* was suggested by Connell and Ladd [12: 2] to the best of our knowledge.

Current intonology accepts the presence of intonation downtrends as a feature of melodic forms in most languages of the world. John Ohala even states that perhaps all languages in the world might exhibit some form of downtrends [23: 31], and Hirst and Di Cristo [16: 18nn] suggest that downtrends need to be described in any comprehensive account of a language intonation structure.

![Figure 1. An illustration of an $F_0$ downtrend captured by a linear regression line. The utterance is part of a continuous narrative and says: “Zopakovala jsem to ještě jednou.” (“I repeated it once again.”)](image)

When studying the past accounts of the downtrends, we noticed that while the terminology or methods of measuring and expressing them differed immensely, quite a few papers were in agreement about the fact that the first person to report downtrends was Kenneth L. Pike in 1945. Our key question in the present paper then asks whether the downtrends were really noticed only as late as in the 1940’s.

## 2 Credits due?

### 2.1 Kenneth L. Pike

The American linguist and polyglot Kenneth L. Pike influenced quite profoundly several areas of linguistics and phonetics. Many authors hint, but some even allege that the first published account of downtrends in intonation was given by Pike in his book *The Intonation of American English* (1945). We will now list the suggestions in that respect and we will discuss their explicitness. The quotes are ordered by the date of their publication.
The phenomenon to which we want to refer, declination, was first reported by Pike

This reference is straightforward and unambiguous. In their paper for an international peer-reviewed journal, the authors do not seem to express any doubt about the fact (and they reassert it on p. 256). Since the paper is designed as an overview of the research area and a discussion of previous interpretations of the phenomenon, the claim sounds quite authoritative and may have been influential.

In languages like Dutch and English, the most important global attribute is the observed tendency of $F_0$ to decrease slowly from beginning to end of an utterance. This phenomenon has already been reported by Pike.

This quote actually represents the same authors as the previous one, only in different order. It can be noted that eight years after the previous claim the wording is less definite here: Pike’s primacy is only suggested. One can only speculate whether this is due to the general shift in academic discourse towards less absolute stipulations or whether the authors’ confidence was weakened by some revelations.

The concept of declination, or gradual falling off of pitch during an utterance, was noted by Pike (1945) and Bolinger (1964) in perceptual data.

The authors do not explicitly claim that Pike was the first, yet they do not refer to anyone before. As the authors are Americans, we could perhaps understand why they name only two American linguists. This kind of myopia is common to all of us.

Pike was the first one who noticed the decrease of $F_0$ through time (drift).

The authors do not seem to hesitate at all – the statement opens the section “Theoretical Elements” in their paper. They do not state where their certainty is coming from, they seem to know. It has to be emphasized here that we do not intend to criticize the authors. Instead, we want to demonstrate how easily all of us are led.

Declination was first reported by Pike: The general tendency of the voice is to begin on a moderate pitch and lower the medium pitch line during the sentence.

The same case as the previous one: there seems to be no doubt on the part of the authors about the claim. The verbatim quote suggests that the impressive Pike’s monograph was at the authors’ disposal, which may have influenced their decision to nominate Pike as the discoverer.

Declination is categorized as one of the global features of the prosodic contour ... and refers to the perceived lowering of fundamental frequency across a spoken utterance (Pike, 1945).”
The form of the reference seems to lead to a single interpretation, but Pike speaks neither of fundamental frequency nor utterance. The authors most probably adapted the idea from a secondary source. It is also possible that they treat their sources with greater freedom than is common in current scientific discourse or that they do not differentiate between pitch and fundamental frequency or sentence and utterance. As to our chief concern at this point, the wording of the reference to Pike does not constitute a strong claim about his ‘prerogative’.


“It has been clear at least since Pike (1945) that $F_0$ tends to decline over the course of phrases and utterances, both in tone languages and in languages like English or Dutch.”

This is an opening sentence of a section in a book printed by a prestigious British publisher. The author quite probably understands the weight vested on the claims in these initial positions, which could perhaps explain why he is careful with the stipulation and leaves the back door open with the expression at least since. Alternatively, he may have already encountered allusions concerning possible earlier accounts. Be that as it may, the book is a very frequently cited source, so its readership is considerably large. Interestingly, the author speaks briefly about downtrends already on p. 16 in introductory remarks, but does not go before the year 1967 in references there [20: 16].

2.2 Indistinct primacy

A large body of research in the area of downtrends has been produced over the last decades. Many authors actually avoid the considerations concerning the first person to report the phenomenon. They nevertheless refer to the past research as required by contemporary reporting conventions, but they stay less explicit. Although such citations are not directly binding, they might be potentially misleading. Our following examples are ordered by the year which is insinuated as initial in downtrend research.

Patterson [26: 25] in his dissertation introduces the concept of declination by reference to Liberman & Pierrehumbert [21], although two pages later he goes back to 1967 to give credit to the IPO researchers, even if they actually introduced the concept of declination two years before that (cf. Cohen & ‘t Hart [8]).

Liberman and Pierrehumbert themselves dedicate major space to downtrends in their extensive and often cited study, yet when it comes to the history of declination research, they do not go further back than to 1969 [21: 163, 224, 232 – note 1].

Ohala, who was known for his vast knowledge of literature even before the era of the internet, produced an ample reference section for his comprehensive overview of tone production, where he cites even an author from the 16th century. However, when he discusses downtrends, he does not go further back than to 1970 [23: 31, 32].

Clark and Yallop produced a best-selling university textbook of the foundations of phonetics and phonology. It is used as a reliable compendium by many scholars in linguistics and related disciplines. However, in their treatment of downtrends they do not go before the year 1975 [7: 284-286].

Equally short ‘memory’ can be also found in the study by Terken [31]. Although declination is a key term in that study, his initial references do not reach beyond the above-mentioned influential study of Liberman & Pierrehumbert from 1984. Judging from his concise style of writing, it is perhaps possible that he liked their simple wording in the definition [31: 1769]. However, in the General Discussion the fact that he belonged to the IPO team is reflected in his reference to the work of his older colleagues in 1975 [31: 1775], referring to ‘t Hart and Collier [32].
A very special case is that of Roger Kingdon [19], whom we could not resist to add to this section. In his widely used monograph, Kingdon introduces the phenomenon quite clearly:

“When a series of level tones occurs in the head and body of a tune, each succeeding level tone is pitched a little below its predecessor; a long sentence consists, therefore, of a slowly descending series of stressed syllables interspersed with unstressed syllables which remain on or very close to the pitch of the stressed syllables...” [19: 25]

On the same page the author emphasizes that slowly descending stressed syllables are “a basic feature of English intonation”. However, he does not refer to any other authors who would corroborate this view. By doing so, Kingdon presents the downtrend as something common and undisputable.

2.3 Earlier accounts

Certain doubts about Pike’s primacy were expressed cautiously by Volín [34: 2], who knew the drawings in the monograph by Armstrong and Ward from 1926 (see below). Only later, he discovered that his doubts were substantiated by the research of Collins and Mees [11], who actually suggested to go even before 1926. This directed us to the oldest source cited in the present paper: Intonation Curves (1909) by Daniel Jones [17]. In this extraordinary work Jones mapped an unprecedented amount of speech material in three languages: English, French and German. All the samples represent acted speech (drama, poetry, acted conversation), and Jones depicts the tunes of utterances with curved line drawings in a musical stave to “combine as far as possible scientific accuracy with practical utility” [17: v]. The curved lines were made through perceptual analysis over repeated playback of short stretches from a gramophone record. The drawings start appearing on p. 3, where there are 10 utterances, of which 8 exhibit a downtrend. An example is displayed in Figure 2.

Collins and Mees in their thorough biography of Daniel Jones refer to hand-written lecture notes in which the English phonetician illustrates downstep of lexical tones in Sechuana (i.e., Tswana) [11: 161]. They document that a couple of years later, Jones presented the downstep in Sechuana in print, specifically in Transactions of the Philological Society [11: 187].

One of the handbooks in long-term high demand was Jones’ Outline of English Phonetics, which was first published in 1918. Its 9th edition was published in 1960 and reprinted several times, also in 1969, two years after the author’s death. Jones kept continuously rewriting parts of the text, improving the transcriptions, and adding materials into the book. The oldest edition that we were able to access was the second one [18]. The chapter on intonation there stretches over 34 pages and Jones uses musical notation together with curved lines to capture the pitch tracks. Many of the examples in the Outline are actually too short to become the domain of a
downtrend, but the longer ones show it unambiguously. More importantly, Rule 10 in the chapter says:

“When in a sentence having a falling intonation there are a number of stressed syllables, the first important stressed syllable generally has the highest tone and the other important syllables form a descending series of notes.” [18: 161]

Harold Palmer is another author who represents what is referred to as the British school or British tradition (e.g., [14: 278] or [14: 301], respectively), and he is also clearly aware of the downtrends. The evidence can be found in his highly praised handbook from 1922 both in some of the diagrams of the tunes there and in the wordings such as: “the first syllable of a superior head is on the high pitch, each successive syllable being pitched slightly lower than the one before” [25: 45]. It is interesting to see that Palmer depicts the downtrend as independent from the final nucleus. Figure 3 presents three different nuclei, each of which is preceded by descending intonation.

![Figure 3. Descending tunes as captured by Palmer [25: 45].](image)

Four years after Palmer and eight years after the first edition of the Outline, two members of the Department of Phonetics at University College London, Armstrong and Ward, published a more extensive and detailed treatise on English intonation [2]. Interestingly, the book is dedicated in capital letters “TO PROFESSOR DANIEL JONES” (Figure 4), but at the same time, Jones gives credits to both authors in his later works and adopts many of their improvements in his own intonological studies. This can be interpreted as an indication of bidirectional influence and gratitude. The book was accompanied by three double-sided gramophone records which comprised Armstrong and Ward’s readings of English passages, and was in print for about 50 years. It inspired a large number of intonologists over the world, and even its critics mostly just repeated what the authors themselves already admitted and defended as their intention in the context of their work.

More to the chief concern of our study, numerous tunes in the book clearly contain undisputable downtrends and the very first example of English intonation on p. 4 is no exception. Right under this first example, there is a sentence: “The stressed syllables form a descending scale.” This assertion is apparently meant as a general rule.

One of the early published reviews of the book was written by Anna Bohnhof [3]. Although the reviewer does not explicitly mention intonation downtrends in any sort of a defining statement, the illustrations of tunes reprinted from Armstrong and Ward (1926) clearly contain them. Their salience is indisputable (Figure 5).
For some reason, the examples are reprinted in the review without the guiding lines demarcating the pitch range so they might be less convenient for reading than the original ones. Nonetheless, 8 out of 16 example tunes have the structural potential to host downtrends, i.e., at least two stressed syllables before the nuclear pitch accent. All of these materialize the potential and display the downtrend. Moreover, one of the tunes demonstrates the so-called declination reset, which is commented upon in the text [3: 270].

Still ten years before Pike, there is another interesting evidence of downtrend awareness clearly documented in the contribution of S. Boyanus to the 2nd International Congress of Phonetic Sciences in 1935 [4]. The Congress took place in London, where Boyanus settled permanently a year before that, obviously after marrying Lilias Armstrong. His presentation concerned principal sentence melodies of Russian – his mother tongue. He used 28 stave illustrations of melodies, of which 8 were long enough and structurally suited for downtrends to unfold in them. Indeed, all eight of them do display downtrends (as an example see Types 1 and 3 in Figure 6).

However, the peculiar point here is that Boyanus never comments upon the downtrend in the conference paper. In a footnote, though, he acknowledges the influence of his wife and her colleague Ida Ward on his work. Therefore, it is quite possible that the clear presence of downtrends in the illustrations was their contribution, while Boyanus himself considered the phenomenon perhaps less important to talk about.
Figure 6. Four examples from Boyanus [4: 111], of which 2 are structurally apt to display downtrends.

3 Discussion

We have demonstrated in our survey that downtrends were noted and reported long before it is commonly acknowledged. Why is the work of D. Jones and his colleagues L. Armstrong and I. Ward not referenced? The hypothesis that it might be unknown can only be applied to Jones’ notes from about 1916 reported in [11]. Both Jones’ Outline and Armstrong and Ward’s Handbook were very well known before the often-cited Pike’s statement from 1945. Could the problem then stem from the fact that Pike’s account is longer, more thorough or more sophisticated? Not at all, since it is not. Pike chooses the label drift for the downtrend and explains it with very simple words in two sentences [27: 77]. He mentions the term briefly in a few occasional remarks later in his book without any additional explanations.

When looking for alternative explanations, we might also speculate about the standing of the American structuralist tradition and the didactically oriented efforts of the London school. Jones and his colleagues never denied their ambitions to help language teachers. As Fox characterizes them: “Their concern is less with theoretically adequate formalization than with the identification of pedagogically useful categories” [14: 127]. Since there used to be, and sometimes still is, certain prejudice against pedagogically oriented research, perhaps writers were shy to refer to sources that might be considered less serious. That would be clearly wrong, and not only because history showed that Jones was in the end extremely influential even without being ‘complicated’.

Nevertheless, the illusions mentioned in the introduction of our paper together with certain carelessness might be the most plausible explanation. The error of present-centeredness of historical accounts was exposed already in 1931 in a brilliant essay by Herbert Butterfield [6], and both historians and philosophers keep elaborating on this serious cognitive bias, knowing that it is omnipresent and difficult to eradicate (e.g., [35]).

In our paper we would like to warn against compulsory historical precision, which would be unattainable, but also against careless referencing to whatever is at hand. Human behaviour shows that despite all the clues we receive from the surrounding world, inclination to extremes is often the prevailing modus operandi. Despite all the resources that are available nowadays thanks to the internet, we cannot obligate authors to know the history of their discipline in every detail. On the other hand, there should be very low tolerance towards haphazard referencing, let alone crediting friends and colleagues for trivial claims of the type “speech is important in our everyday lives”.

We have contributed to the question of historical accounts of intonational downtrends. Apart from the possibility of even earlier observations than reported here, our survey leaves some further questions unanswered. For instance, a thorough research would be needed to find out who actually first appreciated important functions of downtrends in various languages or speaking styles. Similarly, it might be interesting to know who first perceptually tested or at least pointed out the contribution of downtrends to naturalness of the speech signal.
Finally, both experts and laymen may ask whether it really matters who the first discoverer was. The answer would depend on the framework of values in which we choose to operate. If we measure progress by the number of new products that can be bought by consumers, then it does not matter. If, however, we measure progress by our respect for the world, then it does matter. We can only truly admire this world and have faith in it if we understand it without bias.

Acknowledgments
This research was supported by the Charles University project Progres Q10: Language in the shiftings of time, space, and culture.

References


