Meta

ΜΕΤΑ

Detecting meaning through speach

David Shillan

See table of contents

Explore this journal

Cite this article

érudit

DETECTING
MEANING THROUGH SPEECH

A Method, and a Reason, for Tune-Analysis of Language *

Writing in 1937, D. Abercrombie says quite rightly of the gulf separating oral and written communication:

We cannot give a picture of speech by marks on paper; when we say that writing represents speech, we do not use « represent » in the same sense as when we say a picture represents a landscape, but rather as when we say a cross on a map represents a church.[1]

In 1947 the authors of *Visible Speech* say :

Pitch change should be shown by a static rather than a dynamic display. The ear hears pitch as superimposed upon form, and there should be some way found to picture these variables in combination without causing one to detract from the other.[2]

In 1964 a review of « Instrumental Methods for Research in Phonetics », at the 5th International Congress of Phonetic Sciences, says:

The extraction of pitch information is a classic problem in the development of bandwidth compression equipment for voice communication. None of the resulting devices really meets the need for a simple and thoroughly reliable pitch meter of the kind that could be so useful in phonetic research. [This field of research is] an area that has shown few striking gains.[3]

In 1958 it was Gunnar Fant's view [4] that Grützmacher and Lottermoser had been on the right lines in their pitch detector of the early 1930's, but that there had been no progress since then.

What does such a method of recording speech aim at doing ? It is a form of visible speech which concentrates on the rise and fall of the melody rather than on a spectrographic analysis of the components of the articulated syllables.

---

1. D. Abercrombie, *Studies in Phonetics and Linguistics*. Oxford University Press, 1965.
2. R. Potter, G. Kopp and H. Green, *Visible Speech*. New York, Van Nostrand, 1947.
3. F. S. Cooper, « Instrumental Methods for Research in Phonetics ». Fifth International Congress of Phonetic Sciences, 1964.
4. Gunnar Fant, « Modern Instruments and Methods for Acoustic Studies of Speech ». Eighth International Congress of Linguistics, 1958.

Why has it not developed more widely and more rapidly ? Because not much use was seen for it, except perhaps as a visual aid in the teaching of a foreign language.

In fact, the two developments which have taken place have had considerable motivation in exactly this field. At the School of Oriental and African Languages, University of London [5], a pitch detector was developed on a tuning-fork principle, which works reasonably well at least on short utterances; it is certainly useful for demonstrating the tones of tonal language. At the Institut de Phonétique, University of Grenoble, the « Détecteur de mélodie » announced in *L'Onde électrique* [6], with later improvements, has proved its effectiveness for capturing and recording continuous stretches of spoken prose with simultaneous pitch and amplitude information. (A similar apparatus designed at Stockholm has been put on the market at Copenhagen, but my observations refer only to the Grenoble apparatus, which was announced first and which was the one used.)

By recording on Mingograph scrolls the Pitch Detector gives us something like an analog of continuous speech, in the two aspects of melody and volume. The pitch line appears in the upper half of the paper as the lower end of a series of discontinuous strokes downwards from the top of the paper, i.e., lower tones producing longer strokes (slightly closer together). The amplitude appears below this in a series of syllabic bursts of wave activity. (There is some recognizable correspondence of certain wave-forms with certain speech sounds, but the addition of the transcription, by hand, remains a skilled task.) Time is shown by a continuous sine wave, which can be either 10 cm or 20 cm per second.

Now, seeing in our mind's eye this long (theoretically endless) scroll of paper emerging, in real time, from the machine — what use, apart from language teaching, can we imagine for it ? The answer to this depends on the question, too big for full examination here, of the relationships between speech and writing and between listening and reading.

Briefly we can say:

a) « It is an elementary, but often forgotten fact that speech is primary. Men spoke before they wrote, and, indeed, the majority of the population of the world still possess only the oral form of communication with their fellow-creatures. However much independence any form of language may have won for itself, it is in the end only a secondary production, and must ultimately depend for its validity on the spoken word. This should be self-evident, but it is neglected more than one would think by those who discuss the subject. » [7] I quote these words of 1937 because they are surprisingly true still. See also Stetson's striking passage on this subject in *Bases of Phonology* [8]. (Sweet had said in 1877 that all study of language ought to be based on the spoken language.)

5. J. Carnochan, « Pitch, Tone and Intonation ». Fourth International Congress of Phonetic Sciences, 1961.
6. R. Gsell et al., *L'Onde électrique,* May 1963.
7. D. Abercrombie, *Studies in Phonetics and Linguistics.* Oxford University Press, 1965.
8. Raymond Stetson, *Bases of Phonology.* 1945.

b) In the process of silent reading we make use of our experience of speech (both as producers and as receivers) to create some form of kinaesthetic analog which gives us the message. The foreign student who has vocabulary and grammar rules but no speech-experience is often incapable of reading (silently or aloud) so as to catch the author's meaning. A computer can be made to « read » text, but a computer with a dictionary and a syntax-programme would be in the same boat, when it came to « understanding » it, as the foreign student. Even a native reader going through a leading article in *The Times* may grind to a halt until he has fed the necessary speech-form (i.e., a sound form including pitch and stress) into his mind's ear.

Therefore, we say, speech — and so phonetics — is not irrelevant to the treatment of text.

Continuing, again briefly and without examples, we say that a particular form of speech experience which seems not to have become conscious until comparatively recently is that speech (i.e. spoken prose) proceeds in a series of tone-groups which consistently pulse with a binary rhythm, the second stress-point (normally the semantic peak of such an utterance) sometimes being replaced by a « silent beat ». This form may be observed in text. Writers write in a form considerably determined by articulatory factors. This shows as a feature comparable to « phrasing » in music, where a meaningful grouping of notes occurs as a unit independent of the single note, the bar, or the group of bars. So too the « phrasing » in language, spoken or written, is independent of the word, the grammatical phrase or clause, and the sentence — though it may coincide with a syntactic unit. The term « phrasing » was applied to this unit of language in my book *Spoken English* [9].

Now we return to the pitch-detector. The unit of phrasing in English is a group of syllables (average about 8 or 9) within a pitch-period ending in a tonal fall (or occasionally a marked rise, as in many questions). A good deal of fog has been paradoxically created by over-elaborate analysis of stress and pitch [10], while a greatly simplified view allows a consistent form to appear. For English, then, simplifying to no more than 3 basic tones the 40 or more quite correctly analysed by Kingdon, we can with

a raised pitch ˈ
a falling pitch ˋ
and a rising pitch ˊor ͵

identify phrasings with their constituent stress-points (stress in English being a composite function into which pitch enters importantly as well as intensity).

Continuous language can thus be analysed into its phrasings, the only other markers needed being () for the silent beat, and + for a stress-related plurality of words.

---

9. David Shillan, *Spoken English*. London, Longmans, Green, 1954.

10. Roger Kingdon, *The Groundwork of English Intonation*. Toronto, Longmans, Green, 1958. ID., *The Groundwork of English Stress*. Toronto, Longmans, Green, 1958.

Here are two passages, selected by Lascelles Abercrombie [11] as examples of the different nature of rhythm in an emotive and in a conceptual use of language, and « phrased-up » by me to show how the principles of phrasing apply equally to both. It will be observed that the result meets the requirements of both the radio speaker at the microphone and the silent reader in his study.

« O eloquent ()/just and mighty Death !/whom none could advise,/thou hast persuaded;/what none hath dared,/thou hast done;/and whom all the world/hath flattered()/thou only hast cast+out of the world/and despised()/thou hast drawn together/all the far-stretched greatness,/all the pride,/cruelty(),/and ambition of man,/and covered it all+over/with these two narrow words,/Hic jacet.

« Conjunction is the representation/of the synthetic unity/of the manifold()./ This idea of unity, therefore,/cannot arise/out of that conjunction:/much rather does that+idea,/by combining itself/with the representation of the manifold,/ render the conception of conjunction/possible(). »

These forms can be shown by the « Détecteur de mélodie ».

By making the use of phrasings a teachable skill, and by subjecting it to the verification of this machine, the subjective dangers of this method of text analysis can be avoided.

Meanwhile we can find objective and statistical support for long-standing claims (not only Coventry Patmore nearly a century ago [12] but back to Joshua Steele in 1775 [13]) that some form of isochronous rhythm is at work in prose.

Practical applications of this analysis are in fundamental research — where, for instance, we can examine the tendency, observed in many texts and two different languages, for phrasings to be grouped into quatrains or multiples of four; in work such as that of the Terminology Bureau of the High Authority of the European Coal and Steel Community, where a convenient form of context is needed, and in machine translation experiments.

Phrasings in languages other than English call for further research. They can be used in French despite the considerable difference of behaviour in stress between a stress-based and a syllable-based language rhythm.

There is still much to be learned from study of the pitch detector scrolls as visible speech. The extraction of the pitch and other information in digital form awaits programming and other skills but will make phrasing analysis possible over really large areas of text.

DAVID SHILLAN
*Cambridge Language Research Unit
England*

---

11. Lascelles Abercrombie, *The Theory of Poetry*. New York, Harcourt, Brace, 1924.
12. Coventry Patmore, *Amelia, Tamerton Church Tower, etc., with Prefatory Study on English Metrical Law*. London, Bell, 1878.
13. Joshua Steele, quoted in *Studies in Phonetics and Linguistics*. Oxford University Press, 1965.