Europäisches Patentamt
Beschwerdekammern

European Patent Office
Boards of Appeal

Office européen des brevets
Chambres de recours

Veröffentlichung im Amtsblatt ja/Nein
Publication in the Official Journal Yes/No
Publication au Journal Officiel Oui/Non

*001549*

Aktenzeichen / Case Number / N° du recours : T 413/88 - 3.4.1

Anmeldenummer / Filing No / N° de la demande : 84 630 096.0

Veröffentlichungs-Nr. / Publication No / N° de la publication : 0 132 216

Bezeichnung der Erfindung: Signal processing
Title of invention:
Titre de l'invention :

Klassifikation / Classification / Classement : G10L 7/06

## ENTSCHEIDUNG / DECISION

**vom / of / du** 13 April 1989

Anmelder / Applicant / Demandeur : The University of Melbourne

Patentinhaber / Proprietor of the patent /
Titulaire du brevet :

Einsprechender / Opponent / Opposant :

Stichwort / Headword / Référence :

**EPÜ / EPC / CBE** Article 56 EPC

Schlagwort / Keyword / Mot clé : "Inventive step (no)"

**Leitsatz / Headnote / Sommaire**

Europäisches Patentamt | European Patent Office | Office européen des brevets

Beschwerdekammern | Boards of Appeal | Chambres de recours

Case Number : T 413/88 - 3.4.1

# D E C I S I O N
## of the Technical Board of Appeal 3.4.1
## of 13 April 1989

Appellant : THE UNIVERSITY OF MELBOURNE
Grattan Street
Parkville Victoria 3052 (AU)

Representative : Waxweiler, Jean et al
OFFICE DENNEMEYER S.à.r.l.
P.O. Box 1502
L-1015 Luxembourg (LU)

Decision under appeal : Decision of Examining Division 037 of the European Patent Office dated 6 April 1988 refusing European patent application No. 84 630 096.0 pursuant to Article 97(1) EPC

Composition of the Board :

Chairman : K. Lederer
Members : E. Turrini
G. Paterson

## Summary of Facts and Submissions

I. European patent application No. 84 630 096.0 (publication No. 0 132 216) was refused by decision of the Examining Division.

II. The reasons for the decision were that the effective independent Claim 1 as filed on 15 January 1988 failed to meet the requirements of Article 123(2) EPC since, in particular, the description and drawings of the application as originally filed did not include any reference to "both the continuous and discontinuous dominant spectral peaks" of the resulting signal being identified and tracked in time, and that the subject-matter of the claim did not involve an inventive step within the meaning of Article 56 EPC in view of the following documents:

- ELECTRONICS AND COMMUNICATIONS IN JAPAN, Vol. 62-A, No. 4, 1979, pages 10-17, Scripta Publishing Co., Washington, US; S. IMAI et al.: "Spectral Envelope Extraction by Improved Cepstral Method" (D1); and

- THE JOURNAL OF THE ACOUSTICAL SOCIETY OF JAPAN, Vol. 32, No. 1, January 1976, pages 12-23, Tokyo, JP; T. MATSUOKA et al.: "Investigation on Phonemic Information of Static Properties of Local Peaks in the Speech Spectra" (D2).

III. The Appellant lodged an appeal against the decision.

IV. Oral proceedings were held before the Board on 13 April 1989, at the end of which the Appellant requested that the decision under appeal be set aside and that a patent be granted on the basis of Claim 1 filed during the oral proceedings and Claims 2-14 as filed on 11 August 1988, of which Claim 1, the only independent

claim, reads as follows:

"1. A method for extracting word information from a speech signal, comprising the steps of determining whether the speech signal is voiced or unvoiced, and smoothing the voiced elements of said speech signal to facilitate further analysis of the signal, characterized in that the smoothing step comprises removing from or suppressing in said voiced elements of said speech signal the fine structure components of the spectrum caused by pitch harmonics to obtain a smoothed spectrum of the dominant spectral peaks, selecting a limited number of said dominant spectral peaks, and identifying and tracking in time both the continuous and discontinuous dominant spectral peaks of the resulting signal to enable identification of the word information contained in said signal."

V. In support of the patentability of the subject-matter of Claim 1 the Appellant essentially submitted the following arguments:

While it was admitted that a method for extracting word information from a speech signal as defined in the preamble of Claim 1, comprising inter alia the step of determining whether the speech signal is voiced or unvoiced, was actually known in the prior art and disclosed at least implicitly in document D1, which clearly referred to the treatment of either voiced or unvoiced signals, the cited documents could not without hindsight be construed as suggesting to perform also the remaining steps of the claims, in the order indicated therein.

In particular, document D1 did not disclose or suggest any localization and selection of a number of dominant spectral peaks after smoothing of the speech signal.

Document D2 related only to a method for analogically analysing a speech signal by means of a bank of 29 single filters set at fixed central frequencies, in which the speech signal was not submitted to any smoothing procedure.

These documents did not address the further processing of the discontinuous dominant spectral peaks, nor did they even suggest that processing of the discontinuous dominant spectral peaks would be of any interest whatsoever for the identification of the word information contained in the speech signal.

**Reasons for the Decision**

1. The appeal is admissible.

2. The question of the compliance of the application documents with the formal requirements of the European Patent Convention, including Article 123(2) EPC, need not be examined in detail. Even on the assumption that these formal requirements were met, the claims could not be allowed for the reasons indicated below.

3. **Novelty**

3.1 Document D1 discloses a method for extracting word information from a speech signal, which involves smoothing the voiced elements of said speech signal to facilitate further analysis of the signal, as may be seen for example on Figures 7(a), (b), (d) and (e) of page 15, which show both the original speech signal for different voiced sounds and its smoothed spectral envelope as obtained by the so-called "cepstral" method. Since the document refers to the treatment of unvoiced sounds as well (see

for example the Figures 7(c) and (f) of page 15), it at least implicitly calls for some measure for determining whether the speech signal is voiced or unvoiced as further set out in the preamble of Claim 1, which was admitted also by the Appellant at the oral proceedings.

In addition to these features, document D1 also discloses that the smoothing step is performed in such a way as to remove from or suppress in the analysed speech signal the fine structure component corresponding to the periodicity of the sound source (see the paragraph bridging the left-hand and the right-hand columns on page 11), i.e. the pitch component of the spoken sounds (see page 14, left-hand column, last sentence), which in particular comprises pitch "harmonics", thus leaving a smoothed spectrum exhibiting a number of dominant spectral peaks (see Figure 7), as defined in the first portion of the characterising part of Claim 1.

This known method is distinguished from the subject-matter of Claim 1 in that it neither involves selection of a limited number of said dominant spectral peaks, nor identification and tracking in time of both the continuous and discontinuous dominant spectral peaks of the resulting signal, as set out at the end of the characterising part of Claim 1.

3.2 Document D2 discloses a method for extracting word information from a speech signal, wherein a limited number (3) of dominant local spectral peaks as determined by analysing the speech signal by way of a filter bank are selected, identified and tracked in time (every 10 ms.) in order to enable identification of the word information contained in the speech signal (page 13, third paragraph).

In contrast with the subject-matter of Claim 1, this

method does not comprise the step of determining whether the speech signal is voiced or unvoiced, nor does it involve removal from the voiced signal of the fine structure caused by pitch harmonics before selecting a limited number of dominant spectral peaks.

3.3 The remaining documents as cited in the European Search Report do not come closer to the claimed subject-matter.

3.4 For the above reasons, the subject-matter of Claim 1 is considered to be novel within the meaning of Article 54 EPC.

4. **Inventive step**

4.1 The nearest prior art is in the Board's view constituted by the method disclosed in document D1. From speech signals representing either voiced or unvoiced sounds, this method achieves extraction of smoothed signals (spectral envelopes) exhibiting a number of dominant peaks, as represented for instance in Figure 7. Whilst speech signals directly re-synthesized from these smoothed signals are said to be almost undistinguishable from original spoken sounds (page 17, last paragraph), document D1 does not disclose any further treatment to be performed on the smoothed signal to allow identification of the word information.

Accordingly, the technical problem to which the claimed method affords a solution as objectively assessed in view of the nearest prior art is to improve the known method in such a way as to enable also identification of the word information contained in the spoken sounds.

4.2 No contribution to inventive step can be seen in the mere recognition of the above-defined technical problem. For, on the one hand, analysing speech signals in order to

enable identification of their word information content is a common procedure in the art of speech recognition, as evidenced for instance by the disclosure of document D2. On the other hand, the Board cannot see any invention either in the mere idea of contemplating to base a speech recognition procedure on the smoothed signal produced by the method of document D1. For, since this signal is said to allow to synthesize sounds which can hardly be distinguished from the original spoken sounds (last sentences of the summary and of the conclusion), it also clearly contains the word information which is required for speech recognition.

4.3 The skilled person faced with the above defined technical problem would derive from document D2 the teaching that identification of the word information contained in a speech signal can be achieved by selecting a limited number of the dominant spectral peaks, which are identified and tracked in time (the frequencies of three major spectral local peaks picked out in every ten milliseconds from the six largest local peaks of the frequency spectrum are treated as the acoustical parameters which are characteristic of a given utterance; page 13, third paragraph). Indeed, these processing steps are implemented in accordance with the method of document D2 on a frequency spectrum as obtained by submitting the speech signal to a filter bank composed of 29 single peak filters (page 13, lines 12 and 13), without previously smoothing the signal in such a way as to remove therefrom the fine structure components caused by pitch harmonics as disclosed in document D1. However, since for the reasons indicated in connection with the technical problem, the signal produced by the method of document D2 clearly comprises the information content which is required for speech recognition, and since furthermore the filter bank used in the method of document D2 inherently causes some smoothing of the frequency spectrum as well by attenuating the frequency

components which lie between the central frequencies of the individual filters, the skilled person in the Board's view a priori had no reasonable ground to refrain from envisaging to use the method of document D2 to perform speech recognition on the signal derived from the method of D1.

Such use, however, directly results in a method from which the method set out in independent Claim 1 is distinguished only in that the latter explicitly calls for identification and tracking in time of "both the continuous and discontinuous" dominant spectral peaks of the resulting signal. In the absence from document D2 of any instruction to ignore some dominant spectral peaks according to their continuous or discontinuous character,
the Board is convinced that the skilled person would on his own motion at first have considered to exploit all the information made available by the processing method of document D2, namely both the continuous and discontinuous dominant spectral peaks as defined in Claim 1. In addition, if consideration of both the continuous and discontinuous dominant spectral peaks is an essential feature of the method as is implied by the presence of this feature in the claim, the skilled person would not have failed, had he been compelled for any reason whatsoever to ignore either the continuous or the discontinuous dominant spectral peaks, to notice the resulting impairment of the speech recognition capability of the method and accordingly to make use of both types of dominant spectral peaks.

4.4 For the above reasons, the subject-matter of independent Claim 1 is not considered to involve an inventive step within the meaning of Article 56 EPC.

5. Accordingly, Claim 1 is not allowable (Article 52(1) EPC). Claims 2-14 are not allowable either, since they are dependent on unallowable Claim 1.

Therefore, the Appellant's request is rejected.

**Order**

**For these reasons, it is decided that:**

The appeal is dismissed.

The Registrar: The Chairman:

F. Klein K. Lederer