

56 9700660-5

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G10L 5/06	A1	(11) International Publication Number: WO 96/35207 (43) International Publication Date: 7 November 1996 (07.11.96)
(21) International Application Number: PCT/IB96/00396 (22) International Filing Date: 2 May 1996 (02.05.96) (30) Priority Data: 95201139.3 3 May 1995 (03.05.95) EP (34) Countries for which the regional or international application was filed: NL et al. (71) Applicant: PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). (71) Applicant (for SE only): PHILIPS NORDEN AB [SE/SE]; Kottbygatan 5, Kista, S-164 85 Stockholm (SE). (71) Applicant (for DE only): PHILIPS PATENTVERWALTUNG GMBH [DE/DE]; Röntgenstrasse 24, D-22335 Hamburg (DE). (72) Inventors: HÄB-UMBACH, Reinhold; Lemierse Berg 21, D-52074 Aachen (DE). BEYERLEIN, Peter; Ellendorferstrasse 87, D-52078 Aachen (DE). THELEN, Eric; Ursulinerstrasse 17-19, D-52062 Aachen (DE). (74) Agent: STRILAND, Wilfred; Internationaal Octrooibureau B.V., P.O. Box 220, NL-5600 AE Eindhoven (NL).		(81) Designated States: CN, JP, <u>SG</u> /European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: SPEECH RECOGNITION METHODS AND APPARATUS ON THE BASIS OF THE MODELLING OF NEW WORDS (57) Abstract For speech recognition a new word is represented as based on a stored inventory of models of sub-word units. First a plurality of utterances are presented that all should conform to the word. For building a word model from the utterances, these are represented by a sequence of feature vectors. First, the utterances are used to train a whole-word model that is independent of the models of the sub-word units. The length of the whole-word model equals the average length of the utterances. Next, a sequence of Markov states and associated probability densities of acoustic events of the whole-word model is interpreted as a reference template represented by a string of averaged feature vectors. Finally, the string is recognized by matching to models in the inventory and storing a recognition result as a model of the utterances.		