

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 October 2001 (04.10.2001)

PCT

(10) International Publication Number
WO 01/73754 A1

(51) International Patent Classification⁷: **G10L 15/14**,
15/16

PHILLIPS, Michael, S. [US/US]; 653 Concord Avenue,
Belmont, MA 02474 (US).

(21) International Application Number: PCT/US01/09422

(74) Agents: **STEINBERG, Donald, R.** et al.; Hale and Dorr
LLP, 60 State Street, Boston, MA 02109 (US).

(22) International Filing Date: 23 March 2001 (23.03.2001)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data:
09/534,707 24 March 2000 (24.03.2000) US

(71) Applicant (*for all designated States except US*):
SPEECHWORKS INTERNATIONAL, INC. [US/US];
695 Atlantic Avenue, Boston, MA 02111 (US).

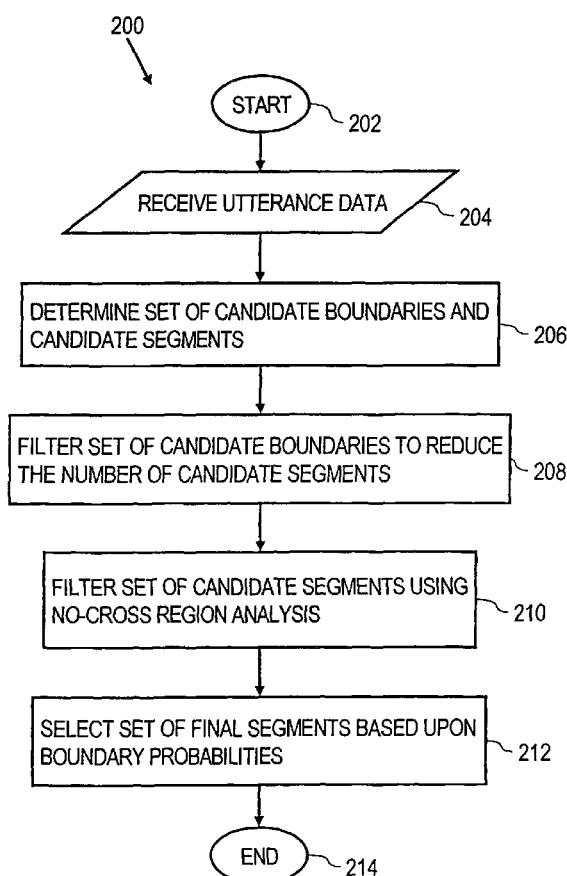
(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **FANTY, Mark**
[US/US]; 108 Union Street, Norfolk, MA 02144 (US).

[Continued on next page]

(54) Title: SEGMENTATION APPROACH FOR SPEECH RECOGNITION SYSTEMS



(57) Abstract: Phonetic units are identified in a body of utterance data according to a novel segmentation approach. A body of received utterance data is processed (204) and a set of candidate phonetic unit boundaries is determined (206) that defines a set of candidate phonetic units. The set of phonetic unit boundaries is determined based upon changes in Cepstral coefficient values, changes in utterance energy, changes in phonetic classification, broad category analysis (retroflex, back vowels, front vowels) and sonorant onset detection. The set of candidate phonetic unit boundaries is filtered by priority and proximity to other candidate phonetic units and by silence regions. The set of candidate phonetic units is filtered using no-cross region analysis (210) to generate a set of filtered candidate phonetic units. No-cross region analysis generally involves discarding candidate phonetic units that completely span an energy up, energy down, dip or broad category type no-cross region. Finally, a set of phonetic units is selected from the set of filtered candidate phonetic units based upon the probabilities of candidate boundaries defining the ends of the unit and within the unit.



Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.