



Control de la rata amb la veu

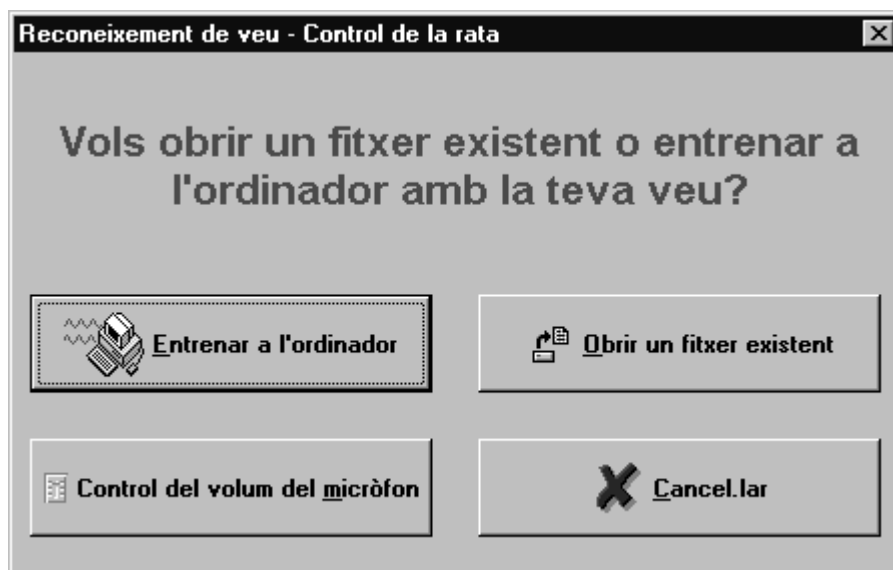
Introducció

La intenció d'aquest programa és tenir el control de la rata amb la veu. És a dir, que amb un micròfon podem controlar els moviments de la rata: esquerra, dreta, puja, baixa, clic.

Està pensat per a persones discapacitades que no poden controlar el *mouse* de forma normal i tinguin control de la veu.

Descripció de la fase d'inici del programa

Quant engeguem el programa apareix després del quadre de crèdit amb el següent quadre:



Aquí tenim quatre opcions:

Entrenar l'ordinador:

Això ens permetrà entrenar a l'ordinador amb la nostra veu. Si fem tot el procés d'entrenament fins al final desapareixerà aquest quadre de diàleg i apareixerà la finestra principal del programa amb el reconeixement activat. O sigui, que si el micròfon recull sons la rata ja es mourà segons les ordres reconegudes. Aprofito l'ocasió per dir que el

tipus de reconeixement és per comparació amb els patrons prèviament entrenats i de moment dóna com a resposta el que més s'hi acosta. Això vol dir que sempre que el micro detecti sons el programa respondrà amb alguna acció. Encara que xiulem.

Obrir un fitxer existent:

Ens permetrà recuperar uns paràmetres de veu prèviament entrenats i desats amb l'opció **Guardar...**

Si escollim un fitxer amb un patró correcte, desapareixerà aquest quadre de diàleg i apareixerà la finestra principal del programa amb el reconeixement activat.

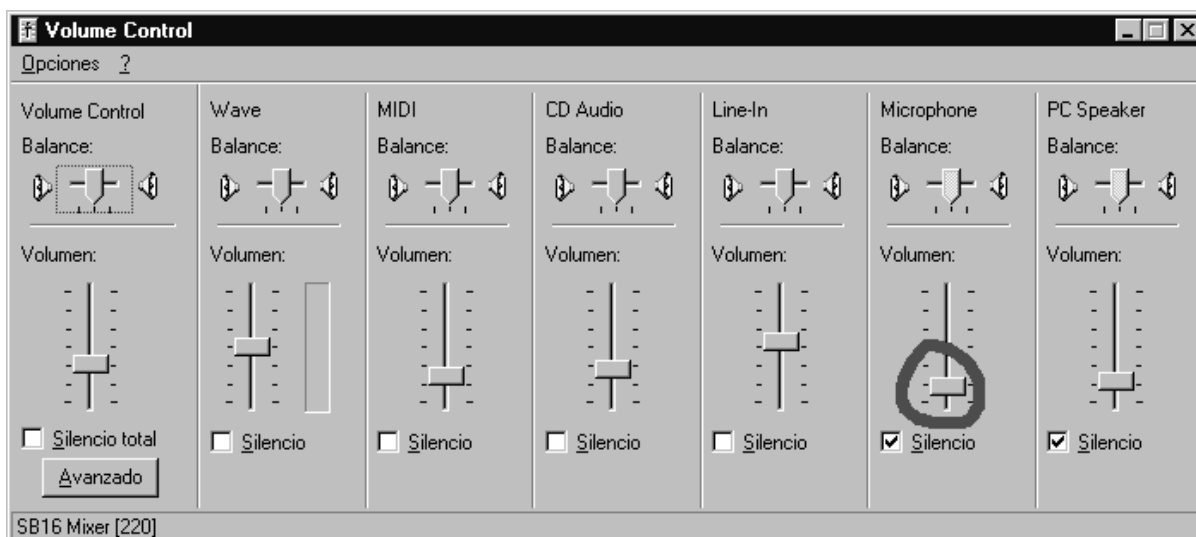
Control del volum del micròfon:

Fa aparèixer una finestra on podrem veure el nivell de so d'entrada quan estigui el sistema de captura de so en funcionament, o sigui, quan estiguem **entrenant a l'ordinador** o quan estem fent el **reconeixement**.



La barra de desplaçament, la fletxa, ens permet triar el nivell mínim de reconeixement, això vol dir que els sons que no arribin al nivell de la fletxa no seran mai reconeguts, els que la passin provocaran el procés de reconeixement. Aquest nivell es pot canviar desplaçant la fletxa esquerra o dreta depenent de si estem en un ambient silenciós o sorollós. També pot interessar modificar això si el nivell de detecció del nostre sistema micròfon-placa és molt o poc sensible.

Aquest quadre també ens permet accedir al programa **Control del volum**, programa que ve amb el Windows o amb la targeta de so, i a partir d'ell pujar o baixar el nivell del volum de l'entrada de so.



Si no es veu el senyal caldrà pujar el nivell del comandament del volum del micròfon. Si, pel contrari, el senyal s'activa i arriba al final molt fàcilment caldrà baixar el comandament del volum del micròfon.

Quant apareix aquest quadre es manté el quadre anterior per si volem fer l'**entrenament** o el **reconeixement**.

Cancel·lar:

Si premem aquest botó no passa res anem directament a la finestra principal del programa i podrem fer les opcions que explicarem a continuació.

Arrencar el programa amb un paràmetre:

Si arrenquem el programa amb un paràmetre, per exemple:

ControlDeLaRataPerVeu Joan.RataTpmSR02

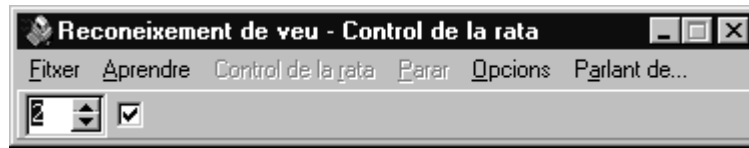
el programa carregarà el fitxer **Joan.RataTpmSR02** i si conté paràmetres correctes de reconeixement passarà a la pantalla principal amb el reconeixement activat.

Recordeu que això es pot fer clicant el botó **Inicio** del Windows i clicant en **executar programa**, on podem escriure la línia de comandes.

També es pot fer en l'explorador del Windows, clicant en el fitxer i arrossegant-lo damunt del fitxer **Control DeLaRataPerVeu.exe**.

Descripció de la finestra principal del programa

La finestra principal del programa té el següent aspecte:



-Hi veiem la barra de títol. Aquí s'escriurà el so o paraula reconeguda.

-Hi ha una barra de menús que descriurem mes endavant.

-Un quadre d'edició que es pot graduar amb fletxetes amunt i avall. És on hi ha escrit el número **2**. Aquest és el número de píxels que es mou la rata. O sigui quant més gran sigui aquest número més de pressa es mourà la rata. Per modificar-ho escrivim el número o cliquem les fletxetes. Tal com és de bon suposar, el número més petit és l'1. Aquest valor el podrem, tal com veurem més endavant, canviar amb la veu.

-Un quadre *checkBox*, per si tenim els botons de la rata intercanviats. Hi ha gent que fa servir la rata amb la mà esquerra i té els botons intercanviats, com és el meu cas, per això ho veieu activat. En principi això no s'hauria de tocar perquè el programa detecta automàticament si hi ha o no els botons intercanviats i o deixa en l'estat *checked* o *Unchecked*, segons hagi detectat. Deixem l'opció a la vista per si algun ordinador té algun comportament estrany. Això serà quant fem l'ordre **clic** i aquesta sigui incorrectament entesa, veurem clic en el títol de la finestra i en lloc de fer **clic** s'obrirà el menú contextual que és el que acostuma a passar quan cliquem amb el botó dret. Quan passi això es tractarà de *check* o *Uncheck*, el botó segons convingui.

-Al costat d'aquest botó hi ha un espai gris. Aquí s'hi escriuran les ordres que pot entendre amb els sons corresponents si el reconeixement es fa per sons. Més endavant entendrem què volem dir amb això.

A continuació passem a descriure les opcions dels menús:

Menú Arxiu

Hi ha quatre opcions, n'hi ha algunes que s'habilitaran o deshabilitaran, es tornen grises, segons l'estat del programa.

Obrir..., permet carregar un fitxer de patrons que prèviament hàgim desat amb aquest programa. Actua de la mateixa manera que ho hem explicat abans. L'extensió per defecte d'aquest programa és: **RataTpmSR02**, tal com sempre es recomana no és bo canviar-la. El programa, automàticament, a l'iniciar-se, s'associa amb aquesta extensió. Si hi hagués un altre programa amb aquesta mateixa extensió, cosa altament improbable, excepte que tinguem dues o més còpies d'aquest programa en el nostre disc dur, ens preguntarà si volem associar-lo a ell. Recordem que si el fitxer conté dades correctes, un

cop carregades ja començarà el reconeixement. En cas que no siguin correctes ja se'ns avisarà mitjançant un quadre de missatge.

Guardar..., permet desar les dades de reconeixement prèviament recollides amb l'opció **Entrenament** en un fitxer del disc. Tal com hem dit abans, l'extensió per defecte és: **RataTpmSR02** i no és convenient canviar-la.

Yeure el nivell del so..., dóna accés al mateix quadre sobre el nivell del volum del so que hem explicat abans.

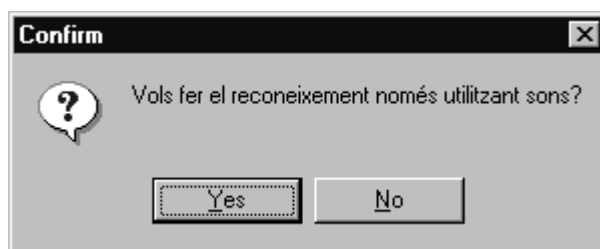
Sortir, permet abandonar el programa.

Menú **Aprendre**

Hi ha només una opció. És així perquè no es cliqui sense voler:

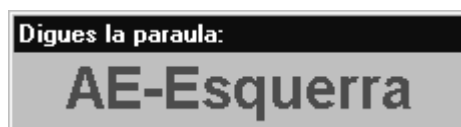
Entrenament..., És l'opció que permet que l'ordinador aprengui a reconèixer la nostra veu. És la mateixa que hem parlat abans en el quadre inicial, però que no hem explicat en profunditat. A partir d'aquí comencem a tractar les particularitats del programa, ja que fins ara, sobre el que hem parlat acostuma trobar-se, coses semblants, a la majoria de programes.

En primer lloc apareix el quadre següent:



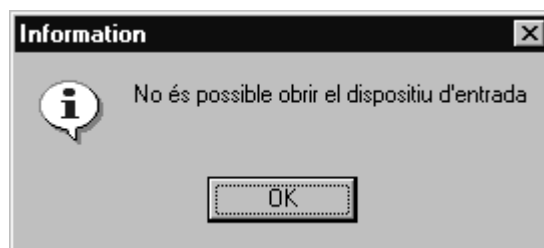
En principi, aquest programa està pensat perquè pugui fer el reconeixement de dues maneres, per paraules o per sons. Cadascuna d'elles té avantatges i inconvenients. El reconeixement per paraules és menys precís, el reconeixement per sons és més efectiu. El reconeixement per paraules és més lent que el reconeixement per sons. El reconeixement per paraules és més fàcil per a l'usuari, sempre està més clar dir *puja* perquè la rata pugui que no un so com *u*.

Bé, la qüestió és que si cliquem retorn farà el reconeixement per sons. Ens apareixerà la següent finestra:



Caldrà acostar-nos al micròfon i dir el so AE (Vocal neutra) durant molt poc temps, un parell o tres dècimes de segon. Quant l'hagi detectat desapareixerà el quadre. Si desapareix massa depressa o sigui, abans que no hàgim dit res potser és per culpa que tenim el micròfon massa sensible. Recordeu el que hem explicat abans sobre utilitzar l'opció control de volum del micròfon. Al principi seria un bon costum tenir aquesta finestra sempre oberta per veure si cridem massa fort o fluix o tenim el valor del micro massa amunt o massa avall. Si, pel contrari, per més que cridem no passa res, o sigui que no desapareix aquest quadre pot ser degut a les següents causes: o el micròfon no està correctament endollat a la targeta de so o el nivell del micròfon està massa avall. En aquest últim cas aneu a l'opció **control del volum**.

En el cas que no hi hagués targeta instal·lada, o estigués mal instal·lada o hi hagués un altre programa fent servir l'entrada de veu com pot ser la **gravadora de Windows** o el **wave estudio** de la **Sound Blaster**. Abans que sortís l'anterior quadre us sortiria el següent:



En aquest cas caldrà que us assabenteu quin dels problemes mencionats és el que falla i corregiu-lo. Fins que no estigui la tarja i el micro correctament instal·lat i que no hi hagi cap programa que la faci servir el programa no pot funcionar. Si l'ordinador tingués més d'una targeta no sé massa que pot passar. Aquest programa fa servir la que l'ordinador hagi configurat com a primera. Però suposo que no es donarà gaires casos que això passi.

Suposant que no hi ha hagut aquesta dificultat i ha aparegut la paraula a reconèixer i l'ordinador l'ha detectat correctament, a continuació apareixerà un nou quadre:



On fem exactament el que hem dit abans podrem entrenar el so que ordenarà a la rata que es mogui cap a la dreta. Si funciona, normalment apareixeran després els següents quadres:

Digues la paraula:
A-Baixa

Digues la paraula:
U-Puja

Digues la paraula:
S-Para

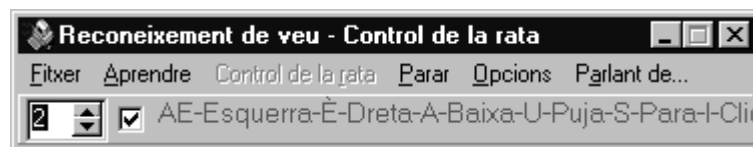
Digues la paraula:
I-Clic

Les dues ordres següents ens permetran controlar la velocitat de la rata, augmentant-la o disminuint-la amb la veu. Quant ho feu anar i funcioni veureu com la rata s'accelera o frena i el quadre d'edició que hem parlat que indicava el número de pixels que es mou la rata va canviant. Direm aquí que la rata fa un salt (moviment) cada dècima de segon aproximadament.

Digues la paraula:
R-ràpid

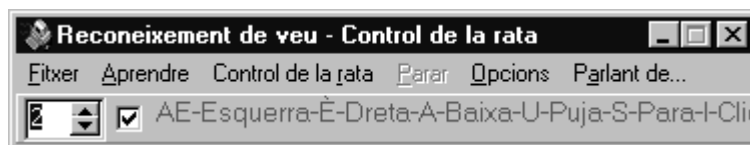
Digues la paraula:
É-lent

Si tot ha ant bé veureu com la finestra principal ha canviat d'aspecte:



Veiem que apareixen les ordres amb els sons, no totes, les que hi càpiguen. Però el que és més important, l'opció **parar**, s'ha activat. Això vol dir dues coses, que el reconeixement ja està funcionant i que si desitgem parar-lo cal que cliquem l'opció **parar**.

Si ho fem veurem que la finestra torna a canviar d'aspecte:



L'únic canvi visible és que l'opció **parar** s'ha desactivat, color gris, i l'opció **Control de la rata** s'ha activat. En aquest moments per molt que cridem el programa no s'escolta. Si volem que ens torni a fer cas, clicarem a l'opció **Control de la rata** i tornarà a començar el reconeixement i interpretació d'ordres.

Quan no funciona el reconeixement veureu que el menú **Guardar com...** està activat, aquesta és l'opció que farem servir si els patrons de reconeixement són prou bons i volem guardar-los per una altra ocasió i no volem fer el reconeixement.

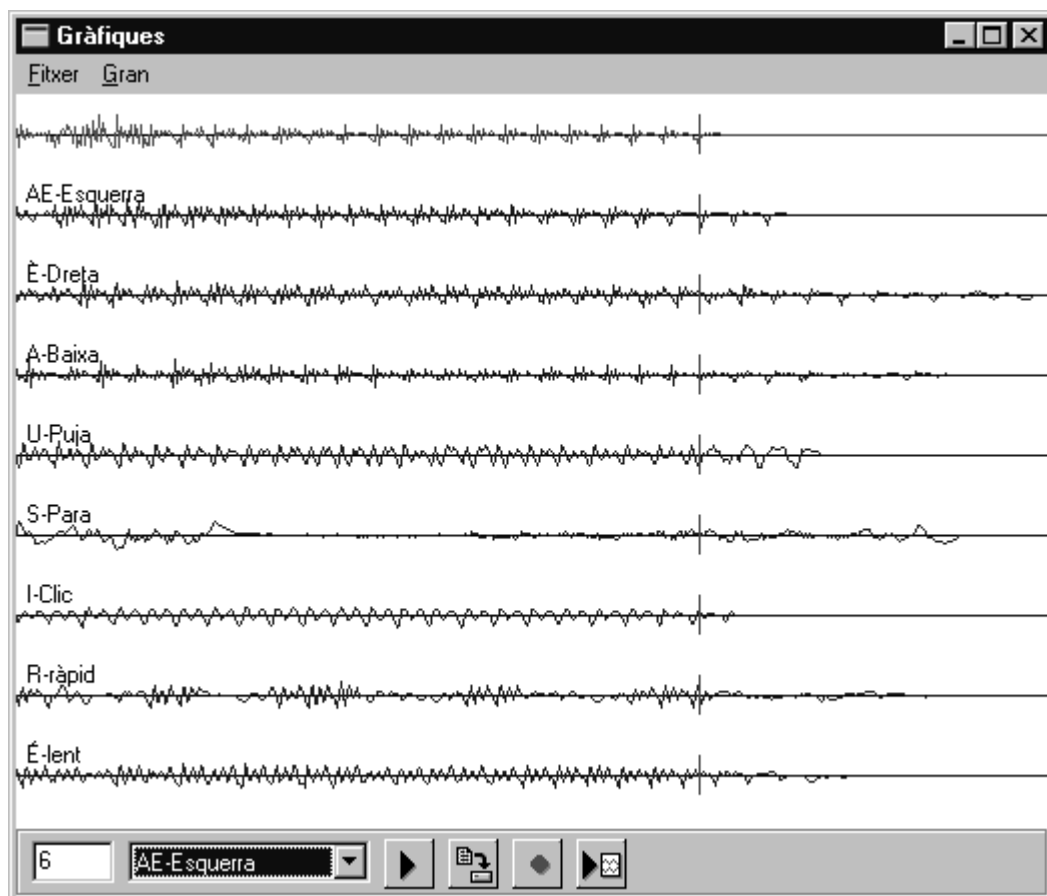
Si fem que el reconeixement estigui actiu, opció **Control de la rata**, podrem fer les nostres primeres experiències de reconeixement automàtic de la veu amb interpretació d'ordres amb el programa del qual estem parlant. Si dieu els sons, AE, È, U, A, I, S, R, É, veureu si el programa us fa cas. És possible que al principi no us cregui gaire, si no es mou cap allà on voleu mireu al títol de la finestra i mireu quins són els sons que detecta.

Què passa si detecta els sons malament

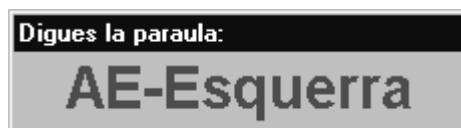
Doneu una mica de confiança al programa, torneu a repetir l'entrenament i torneu-ho a provar. Al principi pot costar una mica, però quan funcioni veureu que el resultat és espectacular.

Hi ha alguna manera de veure perquè falla el reconeixement si és que ho fa? Sí, menú Opcions, opció Veure funcions d'ona.

Si cliqueu aquesta opció us apareixerà la finestra següent:



Aquí tindreu el so que s'ha gravat per fer el reconeixement. Si veieu alguna gràfica de funció que es vegi molt poc potser no s'ha gravat res. Clicant a l'últim botó podrem escoltar tots els sons gravats per si realment hem dit i s'ha gravat el que pensem que hem dit i gravat. Amb el primer botó podem escoltar cadascun dels sons, independentment dels altres, triant-lo en el quadre d'elecció on ara apareix la llegenda *AE-Esquerra*. Si algun so estigués mal gravat podem tornar a fer l'entrenament per aquest últim so clicant en el botó de color vermell que ens farà aparèixer un dels quadres descrits anteriorment i que en aquest cas seria:



I podem tornar a gravar el so. L'altre botó, el de la fletxa assenyalant una espècie de dibuix d'un disc dur serveix per desar la gravació com a fitxer **.wav**. No és una opció que interressi massa per a l'objectiu del programa.

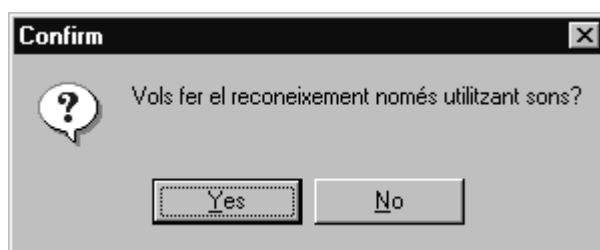
Aprofito per dir que si ens passem per damunt de qualsevol botó apareixerà un *tip* que ens dirà que fa el botó.

El primer so, de color vermell, és al que li fem el reconeixement els blaus sons, els patrons. Si tenim triat aquest primer veureu com al botó de la rodona vermella li desapareix aquesta rodona. És perquè el botó es torna inactiu. Aquest so no és un patró.

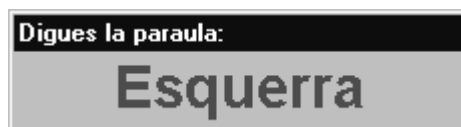
Si alguna persona té certa curiositat i es pregunta si el reconeixement es fa comparant aquestes funcions, us dic que no. Es fa fent unes transformacions matemàtiques a aquestes funcions.

Control de la rata mitjançant paraules

Si quan cliquem a l'opció **entrenament** en el quadre de diàleg:



Cliquem a **No**, en lloc de fer el reconeixement per sons ho farà per paraules



Aquí direm les paraules: **Esquerra, Dreta, Puja, Baixa, Clic, Para, Ràpid, Lent**. Quan vulguem que ens reconegui haurem de dir aquestes paraules. Veureu que el reconeixement és més lent, s'equivoca més, però quan ho fa bé és molt més espectacular.

Final

Control de la rata per Veu és un programa que està contínuament en evolució. És possible que la versió que tingueu tingui més opcions que no estan recollides en aquest manual.

Control de la rata per Veu s'està desenvolupant per en Jordi Lagares i Roset, professor de matemàtiques de l'IES Santa Eugènia de Girona.

Control de la rata per Veu és una primera aproximació al control de l'ordinador, en aquest cas la rata, mitjançant la veu pensada sobretot per a persones discapacitades que no puguin controlar el *mouse* amb la mà i que tinguin un control de la veu.

Tot i les limitacions que pugui tenir aquest programa pensem que pot ajudar a l'escola a cobrir un vuit de manera raonable, reservant altres solucions més costoses per equips més especialitzats.

Per qualsevol suggeriment i aportació, que serà ben rebuda, us podeu adreçar per correu electrònic a: **jlagares@pie.xtec.es**.

Podeu obtenir més informació i altres programes del mateix autor a l'adreça:
[http:// www.xtec.es/~jlagares](http://www.xtec.es/~jlagares).

La versió més actualitzada del present programa la trobareu en aquesta adreça.

Agraïment

L'autor vol expressar a **Joaquim Fonoll** del Programa d'Informàtica Educativa, el més sincer agraïment pel suport, suggeriments fets a aquest programa.