TRANSCRIPCIÓ I AUTENTICACIÓ DE GRAVACIONS EN CONTEXTOS JUDICIALS

Jordi Cicres
Departament de Didàctiques Específiques
Universitat de Girona

1. INTRODUCCIÓ

El llenguatge és l'element omnipresent en totes les societats i té un paper central en tots els àmbits de les relacions socials. També, és clar, en el món de la justícia i del dret. És per això que en les últimes dècades ha estat necessari estudiar els punts en comú que tenen el llenguatge i el dret; aquesta tasca ha estat l'encomanada a la lingüística forense. Aquesta disciplina científica, nascuda a mitjan segle XX, té com a finalitat l'estudi interdisciplinari d'aquesta intersecció principalment des de tres perspectives (Gibbons i Turell 2008): en primer lloc, el llenguatge jurídic (*the language of the law*); en segon lloc, el llenguatge judicial (*the language of the court*), i, finalment, el llenguatge evidencial (*language as evidence*). Aquest darrer punt, probablement el més conegut, consisteix a estudiar el llenguatge en casos en els quals el coneixement lingüístic és necessari per aportar proves (evidències) en una investigació judicial o policial, o bé en un judici.

En aquest article ens proposem tractar dos aspectes relacionats amb el llenguatge evidencial: la transcripció i l'autenticació de gravacions en contextos judicials. Seguint la terminologia de Baldwin i French (1990), en ambdós casos la finalitat d'aquestes tasques tant pot ser investigativa (si l'objectiu del peritatge és ajudar en una investigació) com evidencial (si la finalitat és aportar proves en un judici).

Ambdós casos tenen a veure sobretot amb la fonètica forense (la branca de la lingüística forense que utilitza el coneixement fonètic i fonològic en el context legal); aquesta disciplina s'ocupa també de la comparació forense de veus, la construcció de rodes de reconeixement auditives, la construcció de perfils lingüístics (sexe, edat, nivell educatiu, llengua i dialecte, patologies, etc.) a partir d'una veu i la millora de la qualitat de les gravacions per augmentar-ne la intel·ligibilitat, entre d'altres.

Des del punt de vista legal, la Llei d'enjudiciament civil, en la secció 8 («De la reproducció de la paraula, el so i la imatge i dels instruments que permeten arxivar i conèixer dades rellevants per al procés»), els articles 382 i 383 emmarquen l'ús d'aquestes dues tasques a l'Estat espanyol:

> **Art. 382. Instruments de filmació, gravació i similars. Valor probatori**
> 1. Les parts podran proposar com a mitjà de prova la reproducció davant del tribunal de paraules, imatges i sons captats mitjançant instruments de filmació, gravació i d'altres similars. En proposar aquesta prova, la part podrà acompanyar en el seu cas, transcripció escrita de les paraules contingudes en el suport que es tracti i que resultin rellevants per al cas.
> 2. La part que proposi aquest mitjà com a prova podrà aportar els dictàmens i

© Jordi Cicres (2011)

> mitjans de prova instrumentals que consideri convenients. També les altres parts podran aportar dictàmens i mitjans de prova quan qüestionin l'autenticitat i exactitud d'allò reproduït. […]
>
> **Art. 383. Acta de la reproducció i custòdia dels materials corresponents**
> […] El tribunal podrà acordar mitjançant providència que es realitzi una transcripció literal de les paraules i veus filmades o gravades, sempre que sigui de rellevància per al cas, la qual s'afegirà a l'acta. […]

Són especialment interessants els aspectes subratllats. En primer lloc, la llei preveu que les parts (defensa, acusació, ministeri fiscal) puguin presentar transcripcions que acompanyin (o en ocasions, sobretot en casos poc significatius que substitueixin) l'audició o projecció de gravacions a la sala. En segon lloc, la possibilitat de les parts d'acompanyar aquestes transcripcions (i les gravacions que les contenen) de dictàmens pericials que en demostrin la qualitat i l'autenticitat. I finalment, l'últim fragment subratllat preveu la possibilitat que els jutges puguin demanar la transcripció literal del contingut de les gravacions. Tots aquests aspectes demostren la importància, en els processos judicials, que tenen les proves basades en enregistraments d'àudio, perquè permeten conèixer el contingut exacte de les gravacions.

2. TRANSCRIPCIÓ

El procés de transcripció de gravacions pot semblar a priori una tasca mecànica i a l'abast de qualsevol persona coneixedora de les llengües que s'escolten en els enregistraments. Però la realitat és ben diferent (Fraser 2003). En primer lloc, habitualment es presenten dificultats relacionades amb la inintel·ligibilitat de les veus a causa de sorolls, interferències, presència de diverses veus simultànies, etc. En segon lloc, és convenient d'utilitzar algun sistema estandarditzat de transcripció en el qual quedin reflectits els aspectes rellevants per al cas. En tercer lloc, cal conèixer no només la llengua dels parlants que apareixen en la gravació sinó també els dialectalismes o llenguatge d'argot utilitzat. També, quan en la gravació apareix més d'una veu, el procés de transcripció porta implícita la identificació dels diferents parlants que prenen part en la conversa; aquesta és una tasca que s'escapa a les possibilitats de persones sense una especialització en tècniques de fonètica forense (l'atribució d'una frase a un parlant erroni pot tenir una influència important en el desenvolupament del cas, i caldria que les parts o el jutge encarreguessin un estudi específic d'atribució del fragment *dubitat* [1] a alguna de les persones que prenen part en la conversa). I finalment, en algunes ocasions també és necessari aplicar processos tècnics de millora de la qualitat del so, a fi de millorar-ne la intel·ligibilitat i eliminar-ne sorolls molestos.

De fet, en el sistema jurídic basat en el *Civil Law* —a diferència dels països regits pel sistema del *Common Law*— els procediments legals són eminentment escrits (circumstància, però, que està canviant en alguns països del dret romà) (Gibbons 2003: 27). Per aquest motiu, tots els aspectes relacionats amb el procediment judicial queden fixats en l'escriptura, des dels interrogatoris policials fins al contingut de les gravacions involucrades en un cas.

Però qualsevol transcripció no és res més

[1] En l'argot forense, una gravació dubitada és aquella l'autoria de la qual presenta dubtes o és desconeguda per l'autoritat judicial. Per tant, cal que un expert en fonètica forense la compari amb gravacions *indubitades* (l'autoria de les quals és coneguda sense cap mena de dubte, perquè habitualment s'ha obtingut en presència de la policia o del secretari judicial) dels sospitosos.

© Jordi Cicres (2011)

que una representació imperfecta de l'oralitat, ja que hi ha nombrosos aspectes que es perden durant el procés: el timbre, el to, l'entonació, la intensitat, el tempo, la qualitat vocàlica, les falques dubitatives, etc. I moltes vegades, també, el contingut exacte de les paraules pronunciades. Per això, des de la lingüística i altres disciplines afins (psicolingüística, sociolingüística, adquisició del llenguatge) s'han proposat diferents mètodes de transcripció de l'oralitat que contenen la informació rellevant per als objectius de l'estudi.

Les diferents aproximacions a les metodologies de transcripció (Gibbons 2003) poden anar des de la transcripció fonètica (en la qual es transcriuen els sons produïts pels parlants amb l'alfabet fonètic internacional o alguna adaptació) fins a la transcripció ortogràfica estàndard, passant per opcions intermèdies com les fonèmiques (en les quals només es destaquen aspectes de pronúncia rellevants amb caràcters fonètics, i la resta es transcriuen mitjançant caràcters ortogràfics), la transcripció ortogràfica acurada (incloent les paraules i frases inacabades, les repeticions, les pauses, les falques, etc.) o versions ortogràfiques llegibles en les quals s'eliminen les falques i les repeticions.

En el cas específic de les transcripcions judicials, la norma que s'utilitza en les transcripcions efectuades pels diferents cossos de seguretat a Catalunya i Espanya és transcriure ortogràficament amb versions llegibles, tot i que habitualment es marquen també les falques dubitatives.

El sistema judicial treballa principalment amb les transcripcions de les gravacions per diversos motius. En primer lloc, per estalvi de temps: llegir els passatges rellevants d'una transcripció és més ràpid que escoltar-la. En segon lloc, perquè en facilita l'escolta, en cas que la rellevància del cas faci imprescindible escoltar els enregistraments originals. I finalment, perquè facilita la redacció de la sentència, ja que es pot fer referència a pàgines concretes del sumari.

Ho podem veure en aquest fragment de la sentència [2] sobre el judici pels atemptats terroristes de l'11 de març a Madrid:

> «Por último, las conversaciones de Rabei Osman EL SAYED AHMED en las que, según las acusaciones, se atribuye la autoría intelectual de los atentados al decir que 'el hilo de lo de Madrid fue mio… era mi proyecto más querido', etc., son claramente equívocas —disco compacto unico al folio 14229 y transcrita a los folios 15046, 16747, 24154 y ss.—.
>
> Sobre ellas se practicaron varias pruebas periciales, incluida una conjunta con los intérpretes que hicieron la traducción en Italia.
>
> El día 22 de mayo, el intérprete con número B-12, expuso cómo no oyó las cintas sino que hizo las traducciones sobre las transcripciones en árabe de las conversaciones —ff. 83890 ratificación ante el instructor, 69066 y ss. entrega de la traducción—. Este intérprete aclaró que cuando se habla de grupo no es equivalente a organización y que la expresión "los jóvenes" es muy utilizada y no implica conocer a aquellos a los que se refiere.
>
> El 30 de mayo se practicó una segunda pericial, ésta propuesta por la defensa del procesado, pero practicada por los intérpretes a disposición del Tribunal. Estos, sobre la grabación 1339-12 afirmaron que, al contrario de lo que dicen sus colegas italianos, en la conversación no se menciona a Al Qaeda y no existe la frase de atribución de los atentados de Madrid, concluyendo que

[2] http://www.asociacion11m.org/sumario_20_04/sentenciaFundamentosJuridicosYFallo.pdf

© Jordi Cicres (2011)

> carece de rigor y precisión —ff. 8803 a 8854 del tomo 26 del rollo de Sala—.
> En la tercera y definitiva pericial, en la que estuvieron presentes los intérpretes italianos, todos los peritos españoles, dos de la Unidad Central de Información Exterior, los dos que hicieron la pericial de la defensa y el perito habitual de la Audiencia Nacional, coincidieron en la inexistencia de la frase en la que el procesado se atribuye los atentados. A diferencia de las periciales anteriores, en esta contaron con la grabación depurada o filtrada que trajeron desde Italia los intérpretes de allí. Sin embargo, sus conclusiones no variaron.
> En consecuencia, el Tribunal no puede dar por probada esas conversaciones con el contenido incriminatorio que las acusaciones pretenden.» (pàg. 634-635).

A partir d'aquesta sentència (especialment en allò exposat en el penúltim paràgraf) també podem fer referència a l'anàlisi de fragments disputats (AFD). La diferència d'aquest tipus d'anàlisi respecte de la transcripció del contingut lingüístic és bàsicament que la transcripció és una tasca bàsicament auditiva, mentre que l'AFD implica necessàriament un examen acústic. L'altra diferència principal és el grau de focalització: a diferència de les transcripcions, l'AFD se centra en fragments breus que, per diversos motius, presenten una dificultat específica per poder ser transcrita: o bé hi ha alguns mots inintel·ligibles auditivament, o bé l'atribució de l'autoria del fragment és dubtosa.

Un exemple d'AFD fa referència al cas del judici dels atemptats de Madrid (fragment subratllat de la cita anterior). Inicialment, els transcriptors de les gravacions van incloure-hi referències a Al Qaeda i frases inculpatòries que en realitat no apareixien en els enregistraments originals (com va reflectir també la premsa): «Los traductores afirmaron también que hay omisiones, adiciones, palabras mal interpretadas, sustituciones de contextos y contrasentidos, por lo que concluían que las traducciones realizadas en Italia "no reflejaban fielmente el contenido de las conversaciones grabadas"» (*El País*, 31.5.2007 [3]).

3. Autenticació de gravacions

Una altra de les tasques encarregades als fonetistes forenses té a veure amb l'autenticació de gravacions, és a dir, dictaminar si una gravació ha estat editada (modificada) en algun punt. Aquesta és una tasca compartida habitualment entre pèrits informàtics o enginyers i pèrits lingüistes. Els primers s'encarreguen de desenvolupar l'anàlisi física del suport de gravació. En els casos en què aquest suport és analògic, hi ha tècniques (citades per Bijhold et al. 2007) basades en els efectes magnetoòptics del procés de gravació de les cintes analògiques, tècniques que es basen en els efectes Faraday o Kerr (Boss, Gfroerer i Neoustroev 2003, Boss 2005, Bouten et al. 2007). Mitjançant aquestes tècniques, a partir de l'anàlisi de les cintes amb llum polaritzada, es poden observar —i fotografiar— patrons magnètics que permeten comprovar si una cinta ha estat esborrada o si en algun fragment s'hi ha gravat a sobre. També és possible trobar evidències referents a desajustaments en el procés de gravació que poden permetre identificar la gravadora utilitzada.

En els casos en què les gravacions són digitals, en general cal dir que en l'actualitat no és possible concloure amb una certesa absoluta que una gravació no ha estat

[3] http://www.elpais.com/articulo/espana/Traductores/oficiales/descalifican/prueba/clave/Egipcio/elpepiesp/20070531elpepinac_19/Tes?print=1

© Jordi Cicres (2011)

manipulada, però en canvi —en cas que s'observin senyals d'edició amb les tècniques disponibles— sí que es pot concloure que efectivament hi ha hagut manipulació. El procediment per determinar l'autenticitat d'una gravació digital consta de tres enfocaments diferents: físic, auditiu i acústic.

En qualsevol cas, per poder dur a terme una anàlisi completa és necessari disposar de la gravació original i dels elements de gravació utilitzats (micròfon, gravadora, targeta d'emmagatzematge, etc.). Disposar d'aquests elements és important per determinar si alguns dels indicis de manipulació observats es poden atribuir a la gravadora utilitzada, com per exemple els sistemes de detecció automàtica de veu (segons els quals la gravadora s'activa quan detecta un nivell d'intensitat acústica determinada i s'atura quan el nivell és inferior) o en sistemes d'ajustament del nivell de gravació (que augmenta o disminueix depenent de la intensitat dels senyals acústics detectats).

L'enfocament físic (Grigoras 2005 i 2007) té a veure amb les fluctuacions de la freqüència de la xarxa elèctrica (ENF, amb les sigles en anglès). Aquesta tècnica només es pot aplicar amb gravacions de bona qualitat, ja que consisteix en l'anàlisi del soroll emès per la xarxa elèctrica a la freqüència dels 50 Hz. Per tant, aquestes freqüències són captades per qualsevol aparell enregistrador que estigui connectat a la xarxa elèctrica i que faci ús d'un mètode de gravació i d'un format d'arxiu que realment emmagatzemi aquestes freqüències. En gravacions fetes amb formats comprimits o de baixa qualitat rarament aquestes freqüències queden enregistrades nítidament, i per tant, el mètode és inaplicable.

S'ha demostrat que els patrons de l'ENF varien amb el pas del temps, però són força estables geogràficament. Així, es pot comprovar si una gravació ha estat obtinguda en dos dies diferents (perquè els patrons de l'ENF seran també diferents). A més, si es disposa d'una base de dades de gravacions de les fluctuacions de la xarxa elèctrica, fins i tot pot ser possible determinar en quin moment es va enregistrar cada part de la gravació comparant els patrons de fluctuació de la base de dades amb els patrons de fluctuació presents en cada part de la gravació analitzada.

El segon enfocament, l'auditiu, consisteix en una escolta atenta per part d'un lingüista expert amb l'objectiu de detectar discontinuïtats en l'estructura discursiva o sintàctica, de patrons entonatius, dels patrons del soroll de fons, etc. També és possible detectar sorolls estranys que puguin indicar que s'hagi detingut la gravació i s'hagi represès posteriorment. En aquest enfocament, el coneixement de les estructures lingüístiques per part del pèrit és imprescindible, ja que poden tenir a veure amb qualsevol dels nivells lingüístics: fonètic, fonològic, morfològic, sintàctic, pragmàtic, discursiu, etc.

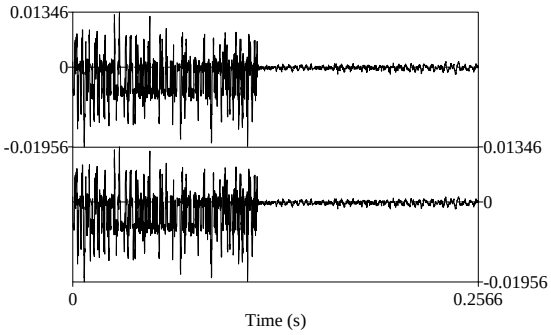
Finalment, l'enfocament acústic consisteix a analitzar l'estructura física (acústica) de les gravacions mitjançant aplicacions informàtiques específiques. Amb aquest programari és possible analitzar diverses representacions gràfiques dels sons. Les més habituals són l'oscil·lograma, l'espectre FFT i l'espectrograma.

L'oscil·lograma representa les variacions d'amplitud al llarg del temps. Per tant, alteracions inexplicables des del punt de vista fonètic o acústic indiquen que la gravació ha estat editada. En la figura 1 s'observa un canvi brusc en la intensitat de l'oscil·lograma enmig d'un fragment que conté només soroll de fons. És difícilment explicable que el soroll de fons hagués disminuït bruscament en l'escenari de la gravació. En canvi, és molt més probable

© Jordi Cicres (2011)

que s'hagi eliminat un fragment de la gravació posteriorment.

Figura 1. Oscil·lograma que mostra un canvi brusc d'intensitat



L'espectre mostra l'amplitud de cada freqüència en un punt temporal o interval temporal determinats. Si s'observen estructures de freqüències inesperades es pot concloure que hi ha hagut algun tipus de manipulació. Aquest és el cas de la figura 2, en la qual s'observa que en el període analitzat no hi ha cap tipus de soroll. Contrasta amb la figura 3, que mostra l'espectre d'una gravació efectuada en un ambient silenciós amb la mateixa gravadora. En aquesta figura 3 s'observa soroll a molt baixa intensitat. Això és una mostra indubtable de manipulació de la gravació, ja que el micròfon utilitzat suposadament en la gravació original podia detectar sense problemes fins i tot sorolls imperceptibles a l'oïda.

Figura 2. Espectre sense cap tipus de soroll

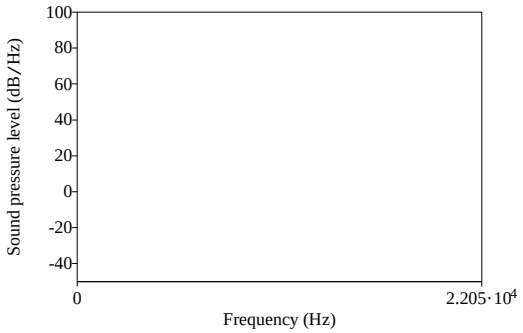
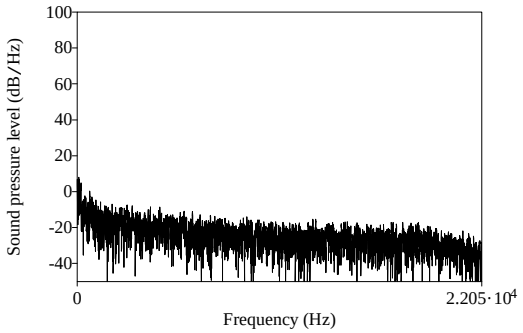


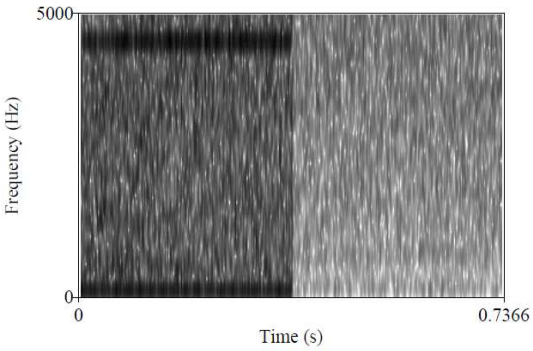
Figura 3. Espectre d'una gravació obtinguda en un ambient silenciós utilitzant la mateixa gravadora que en la gravació de la figura 2



L'espectrograma, per la seva banda, representa les variacions d'amplitud en cada freqüència al llarg del temps. Novament, mitjançant l'espectrograma es poden detectar també inconsistències tant pel que fa a l'estructura espectral com en l'amplitud. Per exemple, l'espectrograma de la figura 4 mostra un canvi brusc en el soroll de fons i la desaparició del soroll present en les franges de freqüències del voltant dels 100 Hz i dels 4500 Hz. Aquests canvis no són explicables si no és que la gravació ha estat manipulada.

© Jordi Cicres (2011)

Figura 4. Espectrograma que mostra un canvi brusc pel que fa als sorolls



Finalment, cal dir també que amb caràcter previ cal analitzar les propietats específiques dels arxius de so. Per exemple, una freqüència de mostreig superior a la que permet obtenir la gravadora suposadament utilitzada per fer l'enregistrament original indica que la gravació probablement ha estat editada en un ordinador o en altres aparells i gravada novament amb unes altres característiques tècniques.

En conclusió, és possible assegurar que una gravació ha estat manipulada quan s'observen evidències de manipulació (en qualsevol dels enfocaments d'anàlisi utilitzats), però en canvi no es pot assegurar amb certesa que no ho hagi estat (ja que és possible que simplement no s'hagin pogut detectar).

4. Referències bibliogràfiques

Baldwin, J. i French, P. (1990). *Forensic Phonetics*. New York: Pinter Publishers.

Bijhold, J., Ruifrok, A., Jessen, M., Geradts, Z., Ehrhardt, S. i Alberink, I. (2007). «Forensic audio and visual evidence 2004-2007: A Review». *15th INTERPOL Forensic Science Symposium*. Lyon, France, octubre 2007. http://www.forensic.to/webhome/enfsidiwg/Interpol-review-2007-audio-visual-evidence-paper.pdf

Boss, D., Gfroerer, S. i Neoustroev, N. (2003). «A new tool for the visualization of magnetic features on audiotapes». *The International Journal of Speech, Language and the Law. Forensic Linguistics*, 10 (2), 255-276.

Boss, D. (2005). «Visualisation of magnetic features by means of crystals. Recent developments and experiences». Estudi presentat a la 14th Annual Conference of the International Association for Forensic Phonetics and Acoustics. Marràqueix.

Bouten, J., Van Rijsbergen, M. i Donkers, S. (2007). «Derivation of a transfer function for imaging polarimetry used in magneto-optical investigations of audio tapes in authenticity investigations». *Journal of the Audio Engineering Society*, 55, 257-265.

Fraser, H. (2003). «Issues in transcription: factors affecting the reliability of transcripts as evidence in legal cases». *The International Journal of Speech, Language and the Law. Forensic Linguistics*, 10 (2), 203-226.

Gibbons, J. (2003). *Forensic Linguistics. An Introduction to Language in the Justice System*. Oxford: Blackwell Publishing.

Gibbons, J. i Turell, M.T. (2008). «Introduction». Gibbons, J. i Turell, M.T. (eds.), *Dimensions of Forensic Linguistics*. Amsterdam/Philadelphia: John Benjamins, 1-6.

Grigoras, C. (2005). «Digital audio recording analysis: the Electric Network Frequency (ENF) Criterion». *The International Journal of Speech, Language and the Law* 12 (1), 63-76.

Grigoras, C. (2007). «Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis». *Forensic Science International*, 167, 136-145.

© Jordi Cicres (2011)