

## L'anàlisi de contingut

### Objectius

En general, l'estudiant ha d'entendre amb precisió què és l'anàlisi de contingut en tant que metodologia específica dins de la qualitativa, i ha de saber utilitzar les seves tècniques més elementals.

En particular ens proposem el següent:

- Saber reconèixer les seves diferents bifurcacions, particularment aquelles que el decanten cap a enfocaments qualitius o quantitius.
- Conèixer bé les diferents fases d'una investigació d'anàlisi de contingut, des del mostratge a la mesura, i ser capaç de resoldre petites qüestions pràctiques en el si de cada una d'aquestes fases.
- Adquirir els recursos necessaris per a dissenyar i planificar, a grans trets, una investigació sencera d'aquesta mena.
- Conèixer el nom, característiques i funcions bàsiques dels principals programes informàtics per a l'anàlisi de contingut, i també els recursos que ofereix Internet en aquest camp.

### Planificació de l'aprenentatge

Mòdul 5. L'anàlisi de contingut					
Títol unitats	Objectius	Nuclis de coneixement	Activitats	Lectures obligatòries	Temps
Presentació de l'anàlisi de contingut	Entendre amb precisió què és l'anàlisi de contingut, en tant que metodologia específica dins de la qualitativa, i saber reconèixer les seves diferents bifurcacions, particularment aquelles que el decanten cap a enfocaments qualitius o quantitius.	Conceptes bàsics  Punts d'origen, direccions i punts d'aplicació de l'anàlisi de contingut		L. Bardin (1986). <i>Anàlisi de contenido</i> . Madrid: Akal. La lectura se centra en el capítol segon: Definición y relación con otras ciencias, pp. 20-35.	4 h
Metodologia de l'anàlisi de contingut	Conèixer bé les diferents fases d'una investigació d'anàlisi de contingut, des del mostratge fins a la mesura, i ser capaç de resoldre petites qüestions pràctiques en el si de cada una d'aquestes fases.	Selecció de material i mostratge  Unitats d'anàlisi  Vies i criteris de categorització  Construcció, estructura i funció d'un sistema de categories AC  La mesura en l'anàlisi de contingut	Exercici de contextualització de mots/frases a partir d'un text de Lévi-Strauss, variant la grandària del context antecedent i consegüent segons el nombre de mots integrants (hi ha un exemple de referència).  Exercici de construcció d'un sistema jeràrquic combinatori de categories AC amb material de comunicació (hi ha un exemple de referència).		8-9 h
Exemples,	Adquirir els recursos	Exemples	Exercici de preses de		4 h

programes, síntesi	necessaris per a dissenyar i planificar, a grans trets, una investigació sencera d'aquesta mena.  Conèixer el nom, característiques i funcions bàsiques dels principals programes informàtics per a l'anàlisi de contingut, com també els recursos que ofereix Internet en aquest camp.	abreujats d'anàlisi de contingut  Tractament informatitzat de dades textuais i de contingut  Programes d'ordinador per a l'anàlisi de contingut  Diagrama d'operacions bàsiques en una anàlisi de contingut	mesura a partir d'un text ja prèviament comentat (sobre el qual s'aplica una nova categorització) i dels exemples del nucli 1 d'aquest mòdul.		
--------------------	---	---	---	--	--

## Presentació de l'anàlisi de contingut

- Conceptes bàsics
- Punts d'origen, direccions i punts d'aplicació de l'anàlisi de contingut

## Conceptes bàsics



**L'anàlisi de contingut és una metodologia, en principi observacional, que permet aïllar unitats, categoritzar, mesurar i establir relacions en corpus de missatges o textos.**



### ■ Generalitats

- En aquest mòdul sintetitzarem les principals línies d'anàlisi que se segueixen a propòsit dels materials de registre descrits fins ara. Les dades obtingudes mitjançant les diverses tècniques de recollida d'informació exposades, en tant que observació participant, no participant o documental, admeten vies també diverses d'anàlisi, dintre –això sí– d'un marc comú.

Aquest marc, recordem-ho, és el que estableix la principal característica de les dades referides, és a dir, el fet que corresponen a un comportament humà amb significat social i cultural, a materials de comportament significatiu i, dintre d'aquest, al que es manifesta pel canal verbal i és codificat i transcrit pel científic, o bé plasmat directament en el paper pel subjecte sota estudi.

#### VEGEU TAMBÉ

Podeu veure la taula de l'apartat "La varietat dels mètodes qualitius: semblances i diferències" del mòdul "Caracterització del marc dels mètodes qualitius en la investigació psicològica i social" d'aquesta assignatura.

En resum, doncs:



**El que es designa genèricament com a anàlisi de contingut és una metodologia destinada, principalment, al tractament i la resolució de dades verbals, transcrits o escrites, tot i que també es pot aplicar a signes i missatges no verbals.**



#### VEGEU TAMBÉ

Podeu veure el subapartat "Precisions conceptuals i terminològiques" i "Què és un document" del mòdul "L'observació documental" d'aquesta assignatura.

Malgrat tot, més enllà de la relativa homogeneïtat dels materials d'estudi, l'anàlisi de contingut inclou diferents

orientacions que, com veurem, són prou contrastades.

A més d'això, convé afegir:



### **En una òptica qualitativa, el significat o significats de les produccions verbals del subjecte constitueixen la via més gran de comprensió per al científic.**



Aquesta metodologia no és pas tan diferent d'altres consagrades a l'observació de camp o resoltes en condicions semicontrolades. Com comprovarem aviat, les fases de la tasca d'anàlisi recorden força les de l'estudi de la conducta o l'acció no verbal mitjançant observació sistemàtica: mostratge, classificació o categorització, mesura, etc.

En un aspecte, però, l'anàlisi de contingut s'aparta de la metodologia observacional estàndard, i és en la seva dependència, més gran o més petita segons els casos, dels enfocaments i les tècniques lingüístiques o semiòtiques, dependència que serà proporcional al grau en què la recerca sigui autènticament qualitativa, és a dir, s'ocupi de signes i significats. Amb aquest prisma, autors com Duverger han considerat l'anàlisi de contingut com un subconjunt de la semàntica.




---

M. Duverger (1975). *Métodos de las Ciencias Sociales* (8a. ed., p. 165). Esplugues de Llobregat (Barcelona): Ariel.

---




---

La nostra presentació de l'anàlisi de contingut ha estat deliberadament esbiaixada cap al pol del significat. Som perfectament conscients que hi ha tendències dintre de l'anàlisi de contingut que defugen o esquiven el significat. Tanmateix, hem volgut oferir una panoràmica d'aquesta metodologia congruent amb l'**enfocament qualitatiu** tal com aquí l'hem entès. Un cop acceptada aquesta exigència, l'anàlisi de contingut ha de recollir obligatòriament la qüestió del significat si no vol perdre la coherència. Això no equival a dir que ha d'excloure les línies de treball compromeses amb els aspectes més formals o quantitativs del text.

Considerem, per exemple, aquestes dues definicions:

- Segons B. Berelson, l'anàlisi de contingut és "l'estudi quantitativ de les característiques manifestes dels missatges".
- Segons P. Stone i O. Holsti, l'anàlisi de contingut és "qualsevol tècnica que permet fer inferències a partir de la identificació objectiva i sistemàtica de certes característiques específiques dels missatges".

Vegeu E. Woodrum (1984). "Mainstreaming" Content Analysis in Social Science: Methodological Advantages, Obstacles and Solutions. *Social Science Research*, 13, 1-19. Noteu que ambdues definicions eviten el terme *significat* i, per contra, se serveixen del terme *missatge*. D'això es pot deduir en clau –justament– d'anàlisi de contingut, que aquests autors estan situats en l'òrbita de la teoria de la comunicació i no en la de la lingüística o la semiòtica.

---

- Així, doncs, el tarannà hermenèutic o interpretatiu del científic qualitatiu es realitza gràcies a l'ús de l'anàlisi de contingut, que és el "banc d'eines" indispensable per tal de dur a terme les interpretacions.

Aquestes interpretacions són bàsicament **contextuals**, arrenquen i descansen en el context en què s'inscriu la unitat textual o d'altra mena que cal analitzar. Més endavant aprofundirem un xic en aquesta noció de context, essencial en determinades orientacions de l'anàlisi de contingut, precisament les que interessin més des d'una òptica qualitativa. Tanmateix, val a dir que aquesta noció ja ha estat prefigurada en un tram anterior de la nostra exposició.

A pesar d'aquesta sintonia de l'anàlisi de contingut amb les interpretacions i –per tant– les inferències, seria un error creure que a cada text o document li pertoca una sola interpretació, la vertadera. Aquesta idea seria més pròpia de l'exegesi dels llibres religiosos o d'aquells que són punt d'origen de doctrines ideològiques o polítiques, ans que d'una mentalitat científica tal com l'hem volgut entendre fins ara.

Certament, de vegades es pensa que el discurs escrit és més fàcil de fixar de manera fiable en la forma i el contingut que el verbal, menys controlable i més susceptible de variació perquè depèn de la memòria individual. Però, tot i que és innegable que un text escrit es presta més fàcilment al control de les seves variacions, també ho és que el sentit sovint depèn de qui el llegeix i de la seva situació com a lector en un món concret.



En relació amb aquest tema, Ian Hodder diu el següent:



"[...] el significat no resideix en un text sinó en la seva escriptura i en la lectura que se'n fa. A mesura que el text és rellegit en diferents contextos rep nous significats, sovint contradictoris però sempre incardinats socialment. Així que no hi ha un significat "vertader" o "original" d'un text més enllà dels contextos històrics específics des dels quals s'interpreta."

Vegeu I. Hodder (1994). *The Interpretation of Documents and Material Culture*. A N. K. Denzin i Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 393-402). Thousand Oaks (Cal.): Sage. (La traducció és nostra.)

Si desitgeu aprofundir-hi més, també podeu consultar D. R. Olson (1998). *El mundo sobre el papel. El impacto de la escritura y la lectura en la estructura del conocimiento*. Barcelona: Gedisa. En aquest volum, la qüestió de la polisèmia dels textos s'explora sistemàticament.

## ■ Materials de l'anàlisi de contingut

A quins materials s'aplica l'anàlisi de contingut? Ja hem respost aquesta pregunta de manera general en la definició de més amunt. Ara, però, la perfilarem millor. Aquesta metodologia, en l'accepció més general, s'aplica al següent:

- **Textos personals espontanis**, categoria en què incloem, a més de les cartes, diaris i altres documents dels quals ja hem parlat, les produccions literàries (novel·la, poesia, teatre, assaig), periodístiques (diaris, revistes), científiques (manuais, llibres i articles teòrics o aplicats) i les peces d'oratória (sermons, discursos parlamentaris, etc.).
- **Arxius, i documents públics o oficials**, com el Codi penal o civil, el redactat de les sentències judicials, els avisos i bans dels ajuntaments, els cartells de la direcció de trànsit, la propaganda electoral o política, les consignes...
- **Textos creats a demanda**, com un relat o narració que se sol·licita al subjecte, o les versions finals que s'hi negocien en una aproximació etnogràfica o etnometodològica, o els autoinformes.
- **Transcripcions de parla fetes pel científic**, tant d'entrevistes, com de dinàmiques grupals, de converses, registrades en viu o en diferit.
- **Transcripcions fetes pel científic de comportaments no verbals comunicatius o expressius** (gest, postura, configuració facial, etc.).
- **Materials no verbals** però estructurats com a signes icònics o simbòlics, en forma narrativa (historieta o còmic, vídeo, cine) o no narrativa (pintura, pintades, fotos, dibuixos, imatges de tanques publicitàries, etc.), que poden incloure, a més a més, una part de text o no incloure-la.

En la resta de la nostra exposició de l'anàlisi de contingut ens centrarem en les quatre primeres categories, corresponents a classificacions de documents ja comentades i que són les que admeten un tractament més estandarditzat dintre d'aquesta metodologia. Els materials inclosos en les dues darreres categories exigeixen molts més recursos lingüístics o semiòtics perquè es puguin estudiar amb un mínim de sentit.



**L'anàlisi de contingut s'aplica a textos o parla transcrita, però també es pot aplicar a altres tipus de signes, no verbals.**

## Punts d'origen, direccions i punts d'aplicació de l'anàlisi de contingut



## L'anàlisi de contingut en perspectiva qualitativa s'orienta cap a la interpretació de significats en context expressiu o comunicatiu.



L'anàlisi de contingut, efectivament, mostra dos vessants ben marcats, un més tècnic i que mira en la direcció dels mètodes més quantitativs i un altre, més interpretatiu, i amb una tradició lingüística o semiòtica al darrere, que mira cap a aquells camps on el treball és més qualitatiu. Aquestes dues tradicions es revelen ja des dels mateixos orígens d'aquesta metodologia i segueixen manifestant-se en la seva posada en pràctica actual.

### ■ Punts d'origen i punts d'aplicació

- Una part important del desenvolupament de l'anàlisi de contingut ha anat indissolublement lligada a les crisis socials i polítiques. El primer ús conegut de tècniques classificables sota aquest epígraf es dona en el marc d'un conflicte religiós a la Suècia del segle XVII. Una col·lecció d'himnes, *Cants de Sió*, era sospitosa d'heretgia i va ser examinada a partir de la freqüència i contextualització de certs símbols religiosos a l'interior dels textos.




---

Vegeu E. Woodrum (1984). "Mainstreaming" Content Analysis in Social Science: Methodological Advantages, Obstacles and Solutions. *Social Science Research*, 13, 3.

---

En el segle XIX els periodistes dels diaris nord-americans ja se servien de recursos com el recompte de temes o paraules per tal d'esbrinar les tendències en la cobertura de les notícies. Ja en el segle XX aquests procediments van ser aplicats al descobriment d'estereotips racials en els llibres de text i en la premsa.

Però l'avanç decisiu d'aquesta línia d'anàlisi de contingut va produir-se en la Segona Guerra Mundial. El govern dels Estats Units va patrocinar un projecte dedicat a l'avaluació de la propaganda política enemiga. Harold Lasswell, encapçalant-lo, va incrementar els recursos metodològics disponibles fins llavors i va obrir-ne el camp d'aplicacions.

Als anys seixanta i setanta del segle XX l'interès es va desplaçar dels textos periodístics o propagandístics a la façana mediàtica de la societat nord-americana, particularment als continguts televisius i a temes candents en aquell moment com la segregació racial o de gènere.

- És fàcil adonar-se que la infància i la joventut de l'anàlisi de contingut, tal com han estat ressenyades fins ara, estan vinculades a la comunicació de masses i a l'univers cultural anglosaxó. Aquest, però, és el vessant d'aquesta classe d'anàlisi menys abocat a l'estudi del significat. Al costat d'aquesta tradició n'hi ha una altra de més europea i més preocupada pel sentit dels textos i pel context que els hi atorga. La seva metodologia (com ja hem dit, molt més deutora de la lingüística) és, fins a cert punt, independent de l'anterior i sorgeix al final del XIX amb voluntat de superar la crítica literària, vista llavors com excessivament arbitrària i personal. Durant el segle XX creix i es desenvolupa al llarg de diferents vies d'aplicació com són la mateixa crítica literària i artística, l'antropologia d'arrel lingüística (dissecció de contes, mites, fàbules) o les interpretacions arqueològiques i històriques.
- Naturalment aquestes dues tendències entren de vegades en contacte, de tal manera que ens podem trobar amb investigacions híbrides entre l'una i l'altra. De fet, la teoria de la comunicació no és, ni molt menys, incompatible amb la lingüística o la teoria dels signes, i viceversa. L'interès o la manca d'interès pel desxiframent de significats i les interpretacions depèn –creiem– de factors ideològics més que teòrics o metodològics.
- En els territoris de la **psicologia** les aplicacions de l'anàlisi de contingut han aportat matisos propis, ja que s'han dedicat a millorar els instruments de mesura de la disciplina. D'entrada, cal dir que l'anàlisi de contingut de respostes verbals s'ha utilitzat, d'una manera o una altra, en la interpretació dels resultats de les **tècniques projectives** des de la seva concepció. Aquesta anàlisi ha impulsat desenvolupaments especials quan ha estat aplicada a proves en què el subjecte havia d'explicar una història a partir d'una escena, com en el test d'apercepció temàtica.

Al marge de les tècniques projectives, el tronc principal d'aplicacions psicològiques de l'anàlisi de contingut se centra en la construcció d'escalas de mesura a partir de les respostes verbals dels subjectes, tant en entrevistes com contestant tests específicament adaptats a aquestes necessitats. L'interès d'aquesta línia de treball rau en què la interpretació i valoració correcta de les respostes, des del punt de vista de la seva significació per al subjecte, il·lumina la relació entre llenguatge i pensament i ajuden a dilucidar el sentit de l'ús d'una paraula o expressió en comptes d'una altra.



Per exemple, els gestos adreçats a objectes es podrien vincular a estats agressius abocats a l'exterior del subjecte, mentre que els gestos de contacte corporal amb el propi cos podrien ser l'expressió d'una violència intrapunitiva o dirigida contra el mateix cos. El contingut dels deliris esquizofrènics varia segons correspongui a una fase aguda o crònica i segons el grau d'integració o autisme.

## ■ Direccions de desenvolupament

Les dues tradicions a què acabem d'al·ludir s'han perpetuat i continuen presents en la recerca actual. Una tipologia de l'anàlisi de contingut les ha de recollir a la força, d'alguna manera. Tanmateix, com ja hem afirmat, aquestes tradicions són compatibles, per la qual cosa els esquemes de classificació que s'hi ha creat al voltant hauran de ser superposables, com anirem assenyalant.

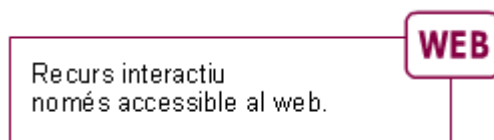
D'antuvi, proposem una classificació de les direccions de desenvolupament de l'anàlisi de contingut basada en l'esquema central de la teoria de la comunicació. Després n'exposarem una de filiació plenament lingüística o semiòtica. Finalment examinarem breument les correspondències de totes dues amb la dicotomia qualitatiu/quantitatiu.

- **Direccions de desenvolupament en el pla de la relació comunicativa.** La teoria de la informació i la comunicació proporciona un fons adequat on situar les principals orientacions de l'anàlisi de contingut.

El fons al qual ens referim ens el subministra l'esquema prou conegut:

**Emissor (codifica, emet) → Missatge → (rep, descodifica) Receptor**

En aquest marc l'anàlisi de contingut es pot dedicar a estudiar el següent:



- **Direccions de desenvolupament en el pla semiòtic o de valor dels signes.** Una divisió clàssica en la semiòtica o teoria dels signes és la que distingeix entre **sintaxi**, **semàntica** i **pragmàtica**.

La sintaxi s'ocupa de la forma i organització en el temps i l'espai dels signes aïllats i de les seqüències o constel·lacions de signes. En certa manera inclou els estudis del **significant**, o sigui, del vehicle o suport físic del signe, com és el so o la traça gràfica d'un mot o la configuració visual d'un senyal de trànsit. Referida als textos, adopta el sentit més popular o escolar i s'interessa per l'ordre en què els mots s'estructuren al llarg del discurs.

La semàntica estudia el **significat** en un sentit: el de la vinculació de la forma del signe (o significant) a objectes o fets de la realitat circumdant tal com són percebuts per l'usuari o la cultura en què habita.

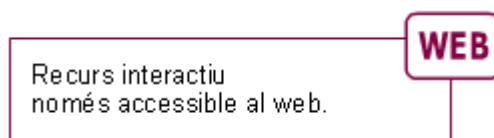
La pragmàtica estudia el **significat** en un altre sentit: el de la vinculació de la forma del signe (o significant) al propi usuari o emissor, al destinatari o a diferents matisos del context o situació social on el signe concret ha estat utilitzat.

Aquests tres vèrtexs del triangle semiòtic estan naturalment lligats entre ells. El significat semàntic completa el pragmàtic, i viceversa; i l'ordre sintàctic dels mots o els signes és inseparable dels significats amb els quals s'associa.

Fetes aquestes precisions indiquem alguns dels camins al llarg dels quals ha progressat l'anàlisi de contingut en cada un d'aquests vessants.



Per tal de fer-ho prendrem com a model l'article de 1995, Anàlisis de contenido, de P. Navarro i C. Díaz, a J. M. Delgado i J. Gutiérrez (Eds.), *Métodos y Técnicas Cualitativas de Investigación en las ciencias sociales* (pp. 177-224). Madrid: Síntesis.



- **Direccions de desenvolupament en el pla de l'oposició qualitatiu/quantitatiu.** Si recordeu el perfil que vàrem traçar dels mètodes qualitatius (perfil al qual ens hem anat referint repetidament) recordareu igualment



que aquests mètodes són eines per treballar la interpretació i el significat. Per consegüent :

#### VEGEU TAMBÉ

Podeu veure el subapartat "Una tipologia dels mètodes qualitius", i especialment la taula que hi ha en aquest subapartat, del mòdul "Caracterització i marc dels mètodes qualitius en la investigació psicològica i social".

WEB

Recurs interactiu  
només accessible al web.



M. Grawitz ens ofereix un exemple prou il·luminador del que volem dir. Enmig de l'eufòrica propaganda nazi, i enmig d'un reguitzell d'èxits dels submarins alemanys, un locutor de ràdio d'aquesta nacionalitat va dir: "No som tan ingenus com per especular sobre el futur basant-nos en aquestes victòries".

Si haguéssim categoritzat, durant mesos o anys, els missatges de ràdio i premsa a l'Alemanya de llavors, tal volta ni tan sols hauríem arribat a confegir una categoria de "missatge no triomfalista" o, si ho haguéssim fet, només li hauríem comptabilitzat **un** cas, el de la frase esmentada.

Des d'un punt de vista freqüencial o estadístic, el que va dir aquest locutor mancava d'importància, no era digne de tenir-se en consideració; en canvi, des d'un punt de vista qualitatiu, la frase es destacava sobre el fons triomfalista de la propaganda i adquiria relleu precisament pel seu caràcter isolat, avisant-nos que potser era ja hora de prevenir els ciutadans sobre futures decepcions.

Vegeu M. Grawitz (1975). *Métodos y técnicas de las ciencias sociales*, (II), 149. Barcelona: Hispano-Europea.



**L'anàlisi de contingut adopta diverses direccions de treball: cap a l'emissor, receptor o missatge; cap a la sintaxi, la semàntica o la pragmàtica; cap a la freqüència o la presència/absència de trets.**

## Metodologia de l'anàlisi de contingut

- Selecció de material i mostratge
- Unitats d'anàlisi
- Vies i criteris de categorització
- Construcció, estructura i funció d'un sistema de categories AC
- La mesura en l'anàlisi de contingut

## Selecció de material i mostratge



**La primera operació en l'anàlisi de contingut consisteix en l'establiment d'un corpus de textos i, dintre d'aquest, en l'elecció dels segments que seran efectivament estudiats.**



Les fases d'una investigació qualitativa coronada amb una anàlisi de contingut són les mateixes, amb els matisos que calgui afegir, que les d'una observació de camp i, en general, que les de qualsevol recerca. Les diferències s'aniran fent paleses i depenen, no solament de les preferències metodològiques de l'enfocament qualitatiu, sinó també del tipus de material analitzat, textos en comptes d'accions o conductes.

Així, doncs, una anàlisi de contingut comença, com no podria ser d'una altra manera, amb una delimitació de l'objecte d'estudi i amb la selecció i fixació dels materials que es pretenen analitzar.

En el llenguatge de l'anàlisi de contingut es diu que aquests materials seleccionats poden haver estat recollits mitjançant la **recopilació** de documents ja existents o mitjançant la **producció** de textos o transcripcions *ad hoc*. Aquesta distinció coincideix íntegrament amb la que nosaltres hem fet entre documents creats espontàniament i documents creats a demanda.

Un cop realitzada aquesta primera selecció de textos podem optar per estudiar-los en totalitat, o bé només limitar la nostra indagació a una part de cada un d'ells. Si preferim aquesta segona opció hem de dur a terme operacions de mostratge sobre el material recollit en primera instància. Els criteris per establir quines parts del text s'analitzen efectivament i quines no poden ser de diversos tipus: estadístics, teòrics, pràctics, normatius, etc.



De la mateixa manera que, en una sessió d'observació, el comportament d'un subjecte no té perquè observar-se i registrar-se d'una manera contínua, en tot moment o en tota unitat temporal de registre, sinó que sovint es comptabilitza només en determinats intervals o punts de temps prèviament fixats, també un text es pot "observar" de manera discontinua, centrant el registre en segments privilegiats i prèviament escollits en funció de diferents tipus de criteris. Amb tot, el paral·lisme entre l'observació del comportament i la de textos té innegables limitacions, de les quals parlarem més endavant.

En definitiva, sembla justificat referir-se a dos tipus d'operacions de mostratge textual que estableixen, al final, les fonts de dades pròpiament dites: per un costat, l'investigador de textos té la possibilitat de fer un **mostratge intertextual** que comporta la selecció dels escrits dels quals s'ocuparà la investigació en curs; per un altre, pot o no dur a terme un **mostratge intratextual** que fixa les parts del textos que seran de veritat registrades, categoritzades, mesurades i analitzades, separant-les de les que es deixaran de banda.

## ■ Mostratge intertextual

- Quan és necessària, la selecció dels documents depèn de l'objecte d'estudi, evidentment, i dels objectius que es marca la recerca en curs. Els interessos i metes de l'investigador delimiten l'**univers** documental on ha de trobar el seu rumb: les cartes d'un polític, els relats dels participants en una crisi familiar, les entrevistes fetes per elaborar una història de vida, etc. I els documents separats i escollits dintre d'aquest univers han de ser **pertinents** o rellevants amb vista a les preguntes que l'investigador es vol respondre.
- El primer que cal fer és establir contacte amb aquest univers documental. Això vol dir que s'han de llegir els textos, probablement no una, sinó diverses vegades. Si els documents han de ser produïts, i no simplement recopilats, la primera feina de l'investigador serà, és clar, construir instruments que generin les respostes adients per tal d'obtenir el material buscat. Els més habituals, com ja sabem, són els protocols d'entrevista.



No és inusual que els documents utilitzats estiguin assenyalats pels propis objectius de la recerca o pels patrocinadors. Aquesta és la situació normal en investigació aplicada. Per exemple, una empresa de perfumeria encarrega a un equip d'analistes la comparació entre els textos de la seva publicitat d'una colònia i els corresponents de la competència. L'encàrrec el propicia una davallada de vendes. En casos així, òbviament, el científic o el professional no poden escollir els materials.

Quan els plantejaments són més amplis el marge de llibertat –i d'incertesa– és més gran. Suposem que ens demanen quins són els eixos de transformació del discurs oficial sobre educació de la joventut durant els darrers anys. Les possibilitats quant al material disponible són enormes: podem recórrer a decrets i notes de premsa del ministeri o la conselleria, a discursos del ministre o del conseller, als discursos en els repartiments de títols o premis, a articles de diaris i revistes, als llibres de text, etc.

- Després de llegir i conèixer els tipus de documents amb què cal treballar ja es pot fer la selecció definitiva sobre la qual es farà l'anàlisi de veritat. Aquesta selecció definitiva donarà lloc a un **corpus**. El corpus, per tant, és el conjunt de materials textuais sobre el qual efectivament treballa el científic social, els escrits que generen la informació que després processa.
- Un cop decidits els criteris selectius que determinen un corpus, el nombre d'elements que conté el corpus ha de ser màxim respecte dels criteris mencionats. En altres paraules: el corpus ha de ser **exhaustiu** en el sentit que cap dels elements previsibles sobre la base dels criteris de selecció no hi pot faltar.
- Parlem finalment d'aquests criteris de selecció. La mostra de textos final ha de ser **representativa** respecte a l'univers del qual prové. Aquesta representativitat apunta al nombre de documents que hi haurà en el corpus i que ha de cobrir la variabilitat que exhibeix en diversos eixos.

En alguns casos, però no en tots, la selecció de materials es pot resoldre mitjançant **criteris estadístics**, els utilitzats en la recerca **nomotètica** o poblacional. Si sabem la distribució dels elements de la mostra respecte a alguna variable podem realitzar un mostratge a l'atzar, estratificat, per quotes, etc.





No seria impossible, posem per cas, conèixer el nombre d'anuncis de cotxes als diaris de Catalunya durant el 2003, classificats per la seva grandària, el nom de la publicació, el mes d'aparició, la marca i model del cotxe, etc., i, tenint en compte totes aquestes fonts de variació, agafar-ne una mostra ajustada a una distribució normal.

Altres vegades els principis que guien la selecció són de caire més **teòric** que estadístic, o estan soldats als mateixos **objectius** de la recerca. Així, si dins d'un epistolari, ens interessen només les missives escrites durant les festes de Nadal, les destinades a certes persones, o les que toquen un tema concret, establim el corpus sobre aquesta base, sense ulteriors complicacions.

Finalment, en un tercer tipus de situació de partida, la representativitat del corpus queda assegurada pel simple fet que univèrs i corpus coincideixen o –el que és el mateix– perquè l'investigador estudia **tots** els documents disponibles relacionats amb el tema que l'ocupa. Aquesta solució és característica de la recerca **idiogràfica** o de l'estudi de casos.

## ■ Mostratge intratextual

- D'antuvi, el text fixat en el corpus, ha de ser **preparat**. Aquesta preparació no és estrictament una operació de mostratge, però la fa possible o la facilita. La preparació ha de ser **material i formal**.

La **preparació material** pretén "posar en net" el text quan no reuneix les condicions de claredat o llegibilitat que se li poden exigir. Una carta escrita amb lletra difícil pot ser mecanografiada; el mateix que un autoinforme o un relat manuscrits. Els documents antics solen demanar una considerable preparació abans de prestar-se al desxiframent i a l'anàlisi. Les transcripcions d'entrevistes o interaccions gravades en magnetòfon o vídeo també es poden incloure dintre de les operacions de "neteja" a les quals ens estem referint.

La **preparació formal** al·ludeix a l'edició dels textos. Així, per comoditat de l'analista, els paràgrafs es poden separar, línia a línia, desglossant les oracions completes en els seus components i numerant aquests; o es poden fixar marges molt amples per a posteriors codificacions; o es pot marcar el text de diverses maneres.

- Pel que fa al mostratge intratextual pròpiament dit el principi orientatiu que cal seguir és el que, un cop delimitat el corpus, com més segments de text n'analitzem, molt més valuosos seran els resultats. De fet, l'ideal és estudiar la totalitat del corpus, atès que un text és semànticament coherent però no homogeni al llarg del seu curs, per la qual cosa la selecció de fragments representatius pot ser problemàtica.
- Si tanmateix cal fer un mostratge intratextual, l'analista haurà de tenir molt present l'estructura del text o el que ja en conegui.

**El text, sens dubte, exhibeix una estructura.** Un discurs presidencial de presa de possessió mostra una organització ritual: hi ha una introducció i una cloenda ben remarcades amb fórmules usualment tradicionals; els afers interns i els econòmics es tracten abans que els internacionals; etc. Una autobiografia sol consagrar més pàgines a la vida adulta que a la infantesa. Un relat palesarà una progressió temàtica molt depenent de la personalitat del seu autor i del tema de què parla: una persona compulsiva introduirà les qüestions centrals de seguida (sobretot si són conflictives), mentre que una de més tranquil·la tal vegada farà un preàmbul llarg.

Per consegüent, serà prudent fer el mostratge dintre de cada sector d'aquesta estructura, que primer que res caldrà conèixer, almenys en una certa mesura. En un discurs, per exemple, caldrà fer una selecció de fragments en la introducció, una altra en el cos central de l'exposició i una altra en la cloenda.

A més, els fragments escollits hauran de tenir una certa unitat: la d'un paràgraf o la d'una secció dintre d'un capítol, posem per cas. En cap circumstància serà convenient separar oracions o mots aïllats, deslligant-los del seu context i fent així la interpretació difícil, si no impossible.

- Semblantment al mostratge intertextual, també en l'intratextual podem recórrer a diferents tàctiques per tal de fer la selecció final de segments de text.

Podem seguir **critèris estadístics**. En aquest tipus de mostratge, tenint en compte que el conjunt de segments seleccionats ha de permetre un repertori i una classificació exhaustius d'unitats de registre i de categories d'anàlisi, l'investigador es pot servir de tècniques habitualment emprades en l'observació del comportament. Efectivament, en aquesta darrera sovint es vol fixar l'abast mostrat amb la finalitat de saber el nombre d'observacions necessàries per tal que el nombre d'unitats de registre i de categories mesurades es pugui considerar complet i no ampliable.

Així, en una anàlisi de contingut, un cop assolida una classificació categorial provisional a partir d'un repertori d'unitats de registre, podem estimar si hem seleccionat prou material dintre del text en qüestió (prou línies, paràgrafs, pàgines) per la fórmula següent:

$$\hat{\theta} = 1 - \frac{N_1}{T}$$

on  $N_1$  és el nombre de categories detectades solament una vegada fins al moment de la decisió i  $T$  el nombre d'unitats de registre recollides també fins aquest mateix moment. Quan l'expressió pren valors propers a la unitat considerem que el nombre de segments seleccionats per a l'anàlisi ja és suficient.

Podem seguir també **critèris teòrics o pràctics**, derivats de la mateixa naturalesa de l'objecte d'estudi o de la demanda que ha motivat la recerca.



Suposem la següent indagació. En els àlbums de còmics clàssics es nota sovint que, originalment, els textos es lliuraven en publicacions periòdiques, pàgina a pàgina o tira a tira. El principal indicatiu d'això en els àlbums publicats complets amb posterioritat es troba en els anomenats efectes de final de pàgina o de final de tira, recursos textuais o visuals amb què els autors pretenien –i aconseguien– mantenir l'interès del lector fins l'endemà, la setmana següent o el dia en què queia el subsegüent lliurament de material. Es parla també, en relació amb aquests recursos, de l'efecte "continuarà".

Doncs bé, imagineu que algú vol estudiar a fons els efectes de final de pàgina. Òbviament, la selecció de parts del text –en aquest cas, vinyetes– obeeirà a aquest objectiu; l'estudi no es farà en tota la superfície de la historieta sinó, com a màxim, en les dues o tres vinyetes finals de cada pàgina. De manera que les operacions de mostratge intratextual no requeriran altra consideració fora de la de quins són els objectius de la recerca.

Vegeu C. Riba (1986). Sobre la utilització de la pàgina en "Tintín". *Neuròptica*, 4, 100-109.

Per acabar, com ja hem dit, l'investigador pot estudiar la **totalitat** dels textos seleccionats, opció practicable si no són molts ni gaire extensos. En qualsevol cas, aquesta és la solució típica de les aproximacions idiogràfiques als textos i no cal puntualitzar que és la favorita de la investigació qualitativa.



**Els criteris de mostratge inter i intratextual poden ser de caire estadístic, teòric o obeir als mateixos objectius de la recerca.**

## Unitats d'anàlisi



**Sobre un corpus textual l'anàlisi de contingut aïlla unitats de registre textuais-contextuals i en fa un inventari.**



Sabem que, indefectiblement, les unitats d'anàlisi que utilitzarà l'anàlisi de contingut es localitzaran al text, seran unitats textuais. Amb tot, caldrà afegir a aquest principi inicial que, atesa l'orientació qualitativa i interpretativa d'aquest tipus d'anàlisi, no solament necessitem unitats de registre en tant que purs "esdeveniments" o peces del text, sinó que també hem d'associar-les amb els seus contextos d'ocurrència, que són els que ens permetran atorgar-les-hi un sentit.

Comptat i debatut, ens faran falta, d'antuvi, dos tipus d'unitats d'anàlisi, els **textuais** i els **contextuals** que, fusionats, donaran lloc a les autèntiques **unitats de registre**. Després, aquestes unitats de registre es classificaran en categories, les variables que sol mesurar l'anàlisi de contingut.

### ■ Unitats textuais

- Les unitats textuais formen el nivell més elemental de referència, i el més molecular, de l'anàlisi. Tanmateix la seva petitesa té, com a límit, la pèrdua de significació. És a dir, almenys en una aproximació qualitativa, la unitat més petita ha d'oferir algun significat, per la qual cosa podem ja avançar-vos que aquest límit estarà en el mot. Per aquesta raó aquestes unitats també s'anomenen de **rellevància** o **significació**.
- Al marge d'això, les característiques i la grandària depenen, com es pot esperar, dels objectius de la recerca, de les mesures que, posteriorment, tenim la intenció d'emprar i de la mena d'inferències que pretenem dur a terme.

- Les unitats textuais han de ser perfectament discernibles en tant que segments d'escriptura i s'han de poder relacionar.
- Les més utilitzades són les següents:

- **El mot**, al qual ja ens hem referit com a peça mínima de significat. Com a unitat es distingeix per ser una forma gràfica (o sonora, abans de la transcripció), simple o composta.

El mot és fàcilment contextualitzable en frases, oracions, etc., i es presta al tractament informatitzat (des de fa anys els ordinadors són capaços de lliurar-nos llistes o recomptes de mots, catàlegs de les frases on figura un mot, etc.).

El problema sorgeix quan un mateix mot, en tant que forma gràfica, ofereix diferents significats o quan hi ha diferents mots que tenen el mateix significat; és a dir, quan hi ha homònims o sinònims. Per això el mot es pot tractar de dues maneres:

- Com a unitat **multívoca**, que presenta diferents valors o significats.
- Com a unitat **unívoca**, que va associada a un sol valor o significat.

El mot "taula" en el segon cas podria proporcionar dues unitats de registre diferents, segons que el seu significat fos "moble amb potes que suporten una superfície horitzontal i que serveix per a menjar-hi, escriure, etc.", o bé "format gràfic de classificació de doble entrada, amb files i columnes". En el primer cas, en canvi, només en proporcionaria una.

- **Frases amb unitat semàntica**, com les frases fetes ("caure molt baix", "donar per fet") i els refranys i dites.
- **Oracions**. De qualsevol tipus, simples, coordinades, subordinades; actives o passives. Els criteris de delimitació de l'oració han de ser, inevitablement, sintàctics o gramaticals. Probablement, després del mot, aquesta és la unitat de registre més emprada.
- **Temes**. Un tema es pot referir a fets, situacions, estats d'ànim; pot ser d'abast individual o col·lectiu. Aquesta unitat, a diferència de l'anterior, és més semàntica que gramatical i, per tant, de delimitació més complicada. En general, però, serà possible fixar els límits del tractament d'un tema en un text.

Podem dir, per exemple, que en una història de vida l'autor comença a parlar de temes centrats en la seva salut a la línia 4 de la pàgina 5 i acaba a la línia 29 de la pàgina 6; torna a mencionar-los a la línia 31 de la pàgina 11 i deixa de parlar-ne a la línia 4 de la pàgina 13; i així successivament.

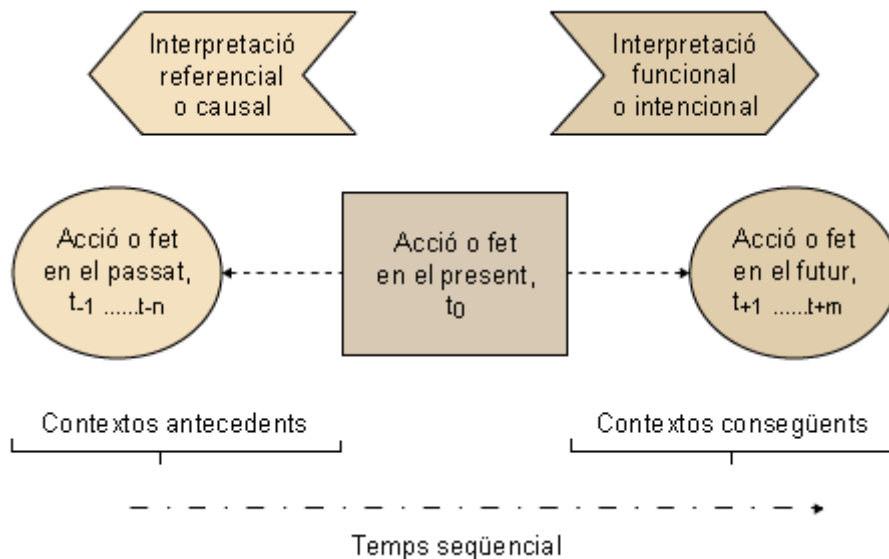
- **Paràgrafs**. Aquesta és una unitat formal, encara que també té connotacions semàntiques, ja que un canvi de paràgraf sol correspondre a un canvi de tema o de matís temàtic. Redueix l'esforç de registre, en ser un segment de text generalment gran. En contrapartida, la tasca de categorització de paràgrafs sol ser més arriscada.
- **Text sencer**. Aquesta solució és practicable si el corpus sencer és escàs, com quan s'analitzen les capçaleres de diaris, pintades o anuncis. Llavors cada text pot ser contemplat com una unitat de registre, comparable amb les altres, i el conjunt del corpus pot veure's com el seu context de referència.

## ■ Unitats contextuais

La unitat contextual o context mínim és el menor segment de text necessari per tal d'interpretar una unitat textual. Les unitats textuais, per si mateixes, poden ser descrites, però no interpretades. Per interpretar-les o per atribuir-los un significat cal posar-les en context, en relació amb l'entorn lingüístic o d'altre tipus que les envolta i els és rellevant. Aquestes relacions poden ser de dues classes, **intratextuals** i **extratextuals**, per la qual cosa distingim dues classes d'unitats contextuais o –simplement– de contextos corresponents a cada una.

- **Contextos intratextuals (o textuais a seques)**

- Són aquells als quals es vinculen les unitats textuais **dintre del mateix text**. Aquestes unitats s'entenen en l'entorn lingüístic, textual, que les conté, i que coincideix amb un segment d'escriptura que les precedeix i amb un altre que les segueix, en conformitat amb la pauta global d'interpretació contextual que ja hem proporcionat en el seu moment.



En una transcripció d'entrevista el context precedent de cada resposta del subjecte pot ser la pregunta prèvia de l'entrevistador, mentre que el consegüent es pot localitzar en la pregunta següent (si es tracta d'una entrevista oberta i en profunditat, que s'adapta sobre la marxa) o ser declarat absent (si es tracta d'un qüestionari rígid, que no pot rebre la influència de l'entrevistat).

- Aquests contextos, per tant, tindran una amplitud variable. El context d'un mot pot ser el mot que el precedeix i el que el segueix; o els dos mots que el precedeixen i els dos que el segueixen; o els tres mots que el precedeixen i els tres que els segueixen; o les frases o oracions que el precedeixen i el segueixen, també en nombre variable. A propòsit del context d'una frase o d'una oració podríem repetir el mateix exercici: una frase podria estar contextualitzada entre la frase antecedent i la consegüent; o entre les dues antecedents i les dues consegüents; etc.

De totes maneres l'amplitud del context està en part determinada per la grandària de la unitat textual considerada. Hi ha una correspondència notòria entre l'escala en què es defineixen les unitats textuales i l'escala en què es defineixen les unitats de context intratextual. El context d'un mot (altres mots, una frase, una oració) tendirà a ser de menor mida que el d'una oració (altres oracions, paràgrafs) i el d'una oració tindrà menys abast que el d'un paràgraf (pàgines, seccions).



Vegeu una petita part d'una llista de contextualitzacions del mot *rights*, generada per ordinador, a partir de documents de les plataformes electorals del partit republicà l'any 1980, amb Reagan de candidat. Aquest és un tipus de contextualització molt corrent anomenada unitat-mot-en-context (key-word-in-context o KWIC). Aquí, més precisament, es tractaria d'una unitat de registre en forma de mot-entre-mots ja que els contextos anteriors i conseqüents estan formats per vuit mots cadascun, tot i que el mot inicial o final pot aparèixer tallat en la impressió.

En la selecció que hem fet del llistat el mot *rights* és una unitat textual unívoca, atès que és portadora d'un sol significat (el corresponent a drets humans, dels ciutadans, etc.).

1)	<i>young people want the opportunity to exercise the</i>	<i>rights</i>	<i>and responsibilities of adults. The republican pa</i>
2)	<i>of their school systems. We will respect the</i>	<i>rights</i>	<i>of state and local authorities in the management</i>
3)	<i>and that new laws with guarantees of individual</i>	<i>rights</i>	<i>are possible and will work. Republicans believe that</i>
4)	<i>it's rethoric. The most flagrant offenders of human</i>	<i>rights</i>	<i>including the Soviet Union will not be ignored</i>
5)	<i>an economic security. Hispanics seek only the full</i>	<i>rights</i>	<i>of citizenship - in education or in law enforcement</i>
6)	<i>t reaffirm our party's historic commitment to equal</i>	<i>rights</i>	<i>and equality for women. We also recall you</i>
7)	<i>xemption from the military draft. We support equal</i>	<i>rights</i>	<i>and equal opportunities for women, without taking t</i>
8)	<i>their emigration is a findamental affront to human</i>	<i>rights</i>	<i>and the decline in exit visas to soviet</i>
9)	<i>policy must be based on the primacy of</i>	<i>rights</i>	<i>and responsibilities. Parental and federal care for e</i>
	.....	....	.....

Font: extret d'R. P. Weber (1990). *Basic content analysis* (p. 45). Newbury Park (Cal.): Sage.

- Les unitats de context, per tant, tenen un estatut relatiu dintre de la jerarquia d'escales possibles. Això permet que, canviant l'escala o el nivell d'anàlisi, una unitat de context intratextual es pugui convertir en una unitat textual. Així l'oració pot ser la unitat contextual del mot però, al seu torn, també es pot emprar com a unitat de registre, enquadrada llavors en el paràgraf –posem per cas– com a unitat de context.
- Aquests nivells o grandàries de context diferents atorguen matisos de significat també diferents a les unitats de text que engloben. No és el mateix assignar significat a un mot per la seva posició dins d'una oració que fer-ho a partir del valor que té en la totalitat d'un escrit. Així, en l'oració "Jo mai deixo de **comprar** el regal d'aniversari per a la meva mare", el valor del mot comprar es fixa principalment en relació amb la destinatària de l'acció (la mare) i la condició de l'acció (l'aniversari). Es tracta d'una compra marcada pels vincles familiars i per afectes o obligacions. Però suposem que aquest mateix mot, dintre de la mateixa oració, s'hagués d'interpretar en el si del diari d'un addicte al consum. Llavors, és clar, la interpretació canviaria i en el context general del document el mot "comprar" adquiriria altres connotacions.
- El context intratextual és el que prefereix l'estructuralisme i el que utilitzen les investigacions lingüístiques i semiòtiques.

#### • Contextos extratextuals (o no textuais)

- Són els contextos als quals es vinculen les unitats textuais **des de dintre a fora del text**. Aquests contextos corresponen a les circumstàncies en què va ser confegit el text, a les condicions de producció. En aquestes condicions es materialitzen els factors pragmàtics que incideixen sobre el text, l'entorn comunicatiu i social, incloent aspectes com la personalitat de l'autor i les vicissituds que patia quan va escriure el document o quan va verbalitzar el discurs que s'analitza; la personalitat o el perfil del destinatari; les motivacions subjacents a la creació del document; el seu rerefons històric; etc.

- Els contextos extratextuals s'adiuen amb **unitats textuais relativament molars o grans**. Això és de sentit comú. En general no sembla gaire prometedora relacionar l'ús d'un mot, per exemple, amb la crisi familiar enmig de la qual algú el va posar sobre el paper; sembla més fructífer lligar aquesta crisi a unitats temàtiques o a documents sencers. Tampoc no semblaria gaire lògic voler interpretar una frase d'un personatge públic en el marc de la seva afiliació a un partit polític. Seria el text sencer el que s'adaptaria millor a un marc tan ampli.
- El context extratextual és el preferit pels estudis sociològics i per la crítica ideològica.

## ■ Les unitats de registre i la seva codificació

- Recapitulant, direm que, normalment, una unitat de registre d'anàlisi de contingut és una unitat de text-en-context, com la de mot-entre-mots que vèiem en l'exemple de més amunt. El context pertinent en el registre és l'intratextual.
- Dit d'una altra manera, la unitat de registre és una unitat textual interpretada o en vies d'interpretació gràcies a la seva contextualització, tant en contextos anteriors, en conseqüents, com en ambdós.
- Aquestes unitats de registre així enteses són codificades mitjançant símbols o notacions i formen la base sobre la qual s'edificarà el futur sistema de categories o variables que permetran quantificar el registre.




---

Considereu aquesta codificació senzilla.

En l'exemple anterior dels textos de les plataformes electorals de Reagan podem vincular el mot *rights* amb diversos marcs de referència: l'emigració, la dona, l'exèrcit, les lleis, la Unió Soviètica i així successivament. Suposem que el vinculem amb mots que denotin igualtat –*equal, equality, etc.*–, interpretant llavors la referència als drets en termes d'igualtat, tant de les dones respecte dels homes, dels emigrants respecte de la població autòctona. Hi ha dues unitats de registre que poden ser codificades o pre-categoritzades així, la 6) i la 7), ja que les esmentades claus contextuais hi són presents, tant en el context anterior com en el conseqüent. En el registre es podrien marcar amb una "Q" (d'*equal* o *equality*).

---




---

Llegiu primerament amb atenció el següent extret d'*El pensament salvatge* (Claude Lévi-Strauss, 1991, Barcelona: Edicions 62; traducció de Miquel Martí Pol), centrat en la relació entre màgia i ciència.

"Això no obstant no regressedem a la tesi vulgar –i d'altra banda admissible en la perspectiva estreta en què se situa– segons la qual la màgia seria una forma tímida i balbucejant de la ciència: car hom es privaria de tot mitjà de comprendre el pensament màgic, si es pretenia de reduir-lo a un moment, o a una etapa, de l'evolució tècnica i científica. Més aviat l'ombra que anticipa el seu cos, és, en un sentit, completa com ell, tan acabada i coherent, en la seva immaterialitat, com l'ésser sòlid precedit només per ella. El pensament màgic no és un principi, un començament, un esbós, la part d'un tot encara no realitzat: forma un sistema ben articulad... En comptes, doncs, d'oposar màgia i ciència, valdria més de posar-los en paral·lel...".

A partir d'aquest text escolliu la unitat de registre "màgia" prenent com a context els nou mots anteriors i posteriors a aquesta; després escriviu els registres resultants. A continuació feu el mateix amb contextos de vuit, set i sis mots, abans i després. Heu de notar que la dificultat d'interpretació o de comprensió creix a mesura que reduïm el marc contextual.

Repetiu el mateix exercici amb la unitat tipus frase "el pensament màgic" prenent ara contextos anteriors i conseqüents de cinc, quatre i tres mots.




---

En els dos casos podeu agafar com a exemple orientatiu el de Weber, exposat en el text. Considerant que el mot "màgia" surt dos cops a l'extret de Lévi-Strauss, lògicament us han de sortir també dues contextualitzacions per a cada un dels marcs contextuais de nou, vuit i set mots que sol·licitem.

En el cas d'"El pensament màgic" us han de sortir igualment dos registres per a cada una de les contextualitzacions de cinc, quatre i tres mots que feu.

---





Les principals unitats textuales, mot i oració, es poden considerar de tipus gramatical o formal; les unitats contextuales poden ser intratextuales i extratextuales.

## Vies i criteris de categorització



L'inventari d'unitats de registre codificades s'ha de classificar en categories de contingut, les variables sobre les quals l'analista farà les seves inferències.



### La categorització d'unitats de registre codificades

- Recordeu que, tal com acabem d'explicar, les unitats de registre que obtenim després del buidatge d'un text són símbols o notacions (lletres, per exemple) que designen les unitats de text-en-context trobades. Així, un registre simbolitzat per "C" ("Comiats"), es pot referir a la unitat "D'acord. Adéu. A reveure", on "Adéu" és el mot interpretat i "D'acord" i "A reveure" els contextos anteriors i conseqüents immediats.

La informació ja codificada d'un registre textual, en tant que recull unitats de text dintre del seu context serà, doncs, heterogènia i complexa, i demana una classificació, implícita ja en el mateix fet d'assignar un símbol a cada una d'aquestes unitats de text-en-context. Sobre aquesta diversitat o riquesa l'analista desproveït de recursos difícilment podria definir i mesurar variables, ni fer inferències sense risc d'equivocar-se. Per tal de salvar aquest obstacle la informació en qüestió s'ha de classificar a partir de les similituds i les diferències. En definitiva, s'ha de categoritzar en classes lògiques, semàntiques, pragmàtiques o sintàctiques.

Aquesta categorització equival a una condensació del material i a una reducció sistemàtica de dades. Així perdem indubtablement informació de detall, a petita escala, però, en rescabament, guanyem capacitat de fer generalitzacions a més gran escala, acomodant la informació recollida en una espècie de taula amb les files i columnes.

- Salta a la vista que, del que es tracta, és de construir un sistema de categories formalment idèntic als que ja hem vist en un altre lloc. No obstant, els materials verbals tenen una naturalesa substancialment diferent a la dels no verbals. Són materials que ja estan interpretats pel seu emissor o, també, pel seu receptor, abans que el científic n'assagi la interpretació.

← Una acció **no verbal** ha de ser definida verbalment, en un codi i en un medi que no són els que li pertocuen; no són els que modelen originalment el moviment conductual.

→ Per contra, una **acció verbal** o una frase ja presenten una definició o caracterització o, almenys, **són definibles en el mateix llenguatge on foren pronunciades o escrites**.

- Tot això, al costat d'altres factors que anirem citant, justifica un tractament específic del tema de la categorització de textos, materials que oposen dificultats específiques a la segmentació i classificació. Parlarem, a fi de no ser gaire redundants, de categorització AC per tal de referir-nos a la creació de categories en l'anàlisi de contingut.

### Punts de partida i criteris per a la categorització AC

- Ens trobem amb tres **punts de partida** ben diferenciats:

→ **Existeix ja un sistema de categories o classificació adequat al material que es vol analitzar.**

La seva aplicació correspon, per tant, a una estratègia *up-down* o hipotètico-deductiva. El sistema ha sorgit d'estudis previs sobre materials semblants als que ara es volen categoritzar. De la mateixa manera que no cal construir un test nou cada vegada que es vol mesurar un tret de personalitat, en el benentès que ja n'hi ha de disponibles –i d'idonis– en el mercat, tampoc caldrà bastir un sistema de categories quan ja n'hi ha d'adients a l'abast de l'investigador.

Sovint aquests sistemes ja consagrats han rebut una certa estandardització, avantatge que els torna molt rendibles quan es desitja fer comparacions entre els resultats de diferents investigacions que se n'han servit.

Avui dia aquests protocols estandarditzats solen estar implementats en programes informàtics i llavors es denominen **diccionaris o índexs**. Un diccionari categorial està organitzat com un tesaurus: les entrades es concentren en conceptes clau i, dins de cada una, hi apareixen els mots o expressions relacionats d'alguna manera amb els conceptes

← **No existeix un sistema de categories adequat al material codificat.**

Aleshores l'analista ha de construir-lo, col·locant-se indefectiblement en una perspectiva analítico-inductiva.

Del procediment de construcció, semblant al següent en el cas de les categoritzacions conductuals, ens n'ocuparem a bastament aviat.

↔ **Una part del text ja està categoritzada, mentre que l'altra no.**

El cas més corrent és el de la transcripció d'una entrevista.

Per una banda les preguntes de l'entrevistador poden veure's com categories de la seva activitat verbal, sobretot si formen part d'un qüestionari preparat. Tocant a aquestes preguntes, doncs, estaríem en la primera situació de partida.

Per altra banda, les respostes de l'entrevistat han de ser encara contextualitzades, codificades i –finalment– classificades. Així, tocant a les respostes estaríem com en la segona situació de partida, és a dir, sense sistema de categories.

Ara bé, si les preguntes corresponguessin a una entrevista oberta, en la qual no estan fixades prèviament, ens acostaríem de nou a la segona situació de partida, aquella en què tota la feina està per fer.

- Els **criteris de classificació** utilitzats per a categoritzar un text vindran dictats, com sol ocórrer, pels objectius de l'anàlisi, però també per l'entorn teòric en què es mou el científic.

En termes generals aquests criteris poden considerar-se semiòtics o comunicacionals, malgrat que aquesta adscripció no sempre es faci explícita. Per tant, encaixen en el ventall d'orientacions de l'anàlisi de contingut ja enumerades abans.

A continuació, però, per tal de no estendre'ns excessivament, encongiem aquest ventall restringint la tipologia de criteris als **semiòtics** (sintàctics, semàntics i pragmàtics) i prescindint dels comunicacionals. A fi de compensar aquesta restricció donarem als tipus finalment utilitzats una accepció àmplia i procurarem subsumir-hi els criteris de caire comunicacional (emissor, receptor, missatge, etc.).

- **Sintàctics:** en el pla comunicacional correspondrien a la forma o estructura del missatge.

Aquí trobem les distincions **gramaticals** o **sintàctiques** estrictes (subjecte/predicat/complement/...; nom/adjectiu/verb/adverbi/...). També cauen en aquest epígraf les categories anomenades "**de forma**", referides més aviat a aspectes **retòrics o estilístics**. Lasswell, per exemple, distingia entre declaracions d'identificació ("Jo sóc comunista", de preferència ("Els comunistes tenen raó en voler fer la revolució") i de fet ("Els comunistes conqueriran el món"). En termes retòrics podem diferenciar argumentacions que recorren a autoritats d'argumentacions que fan valer la lògica.

- **Semàntics:** en el pla comunicacional correspondrien al contingut del missatge.

Aquí s'inclouen les classificacions basades en **temes i en conceptes**, i també les de tipus **lèxic** (agrupacions de sinònims, homònims).

En la literatura, aquests criteris s'apliquen a categories anomenades **de matèria**, que contempnen temes de conversa (la moda, l'esport, la política, homes enfront de dones, la salut, els diners); temes de programes radiofònics (noticiaris, xafarderies, música, tertúlies, magazins); televisius ("culebrotos", telediaris, pel·lícules, concursos, tele-escombreries); etc.

Igualment podem encabir en aquest apartat les categories dites d'**apreciació** o de **valors**, les quals engloben dimensions com "aprova-desaprova", "afirma-nega", "bé-mal", i així successivament.

- **Pragmàtics:** en el pla comunicacional es referirien a l'emissor, a la situació d'emissió, al canal de transmissió, al receptor, als efectes copsats en aquest darrer.

Entre aquests criteris hi comptaríem els que afloren quan hi ha diferents utilitzacions del mateix text en diferents contextos (per exemple, diferents usos d'un mateix missatge segons els propòsits de l'emissor o el seu punt de vista).

Les tàctiques seguides per l'autor del text per tal d'aconseguir els efectes buscats també donen lloc a categoritzacions pragmàtiques, com "propaganda", "negociació", "exhibició de poder", "seducció", "suborn", "adulació", etc.

També incloem aquí les classificacions fetes sobre la base de la intensitat dels efectes en el receptor. En aquest sentit hi haurà textos o missatges més potents que altres.

Finalment, formen part igualment dels criteris pragmàtics aquells que resideixen en els trets diferencials dels actors o protagonistes del text (els personatges de la narració) i els que permeten establir perfils d'emissors o de fonts de procedència, com també de receptors.



Per tal de completar els exemples amb què hem anat il·lustrant els conceptes anteriors, n'afegirem alguns més presentant-los ja amb el format d'inventari d'ítems o elements. Aquests exemples ens facilitaran la preparació del terreny pel següent pas en la nostra exposició: la construcció d'un sistema de categories AC.

En l'àmbit dels estudis clàssics trobem la classificació proposada per **H. D. Lasswell per a l'anàlisi dels fenòmens polítics** en els textos. La llista fa gala d'una clara orientació pragmàtica.

1. Actors o persones dedicats a la política ("polítics" en clau popular).
2. Comunitats de referència que figuren als textos (americans, jueus, comunistes, partits, sindicats).
3. Organitzacions presents als textos (Congrés, Parlament, Senat, Tribunal Suprem).
4. Estats o situacions de la política (guerra/pau; crisis, reformes).
5. Ideologies o sistemes polítics (democràcia, dictadura, liberalisme).

En els estudis –també clàssics– de Stewart sobre la premsa americana es distingeix d'antuvi entre fets i comentaris i, dintre de cada una d'aquestes classes, se separen diferents subcategories:

1. **Fets**
  - 1.1. Fets presentats com a certs
  - 1.2. Fets presentats com a possibles
  - 1.3. Fets la certesa dels quals es valora per un judici
  - 1.4. Fets la incertesa dels quals és objecte de valoració
2. **Comentaris**
  - 2.1. Personals de l'autor o de la font directa de procedència
  - 2.2. Indirectes d'emissor o font determinats
  - 2.3. Indirectes d'emissor o font no determinats
  - 2.4. Cita directa de font determinada
  - 2.5. Cita directa de font no determinada

Vegeu M. Duverger (1975). *Métodos de las Ciencias Sociales* (8a ed., pp. 182-183). Esplugues de Llobregat (Barcelona): Ariel.

Pareu atenció ara en aquesta classificació de Berelson (un altre pioner de l'anàlisi de contingut) i Salter. Ordena els objectius vitals que posen de manifest les telenovel·les i revistes populars.

1. **Objectius afectius o "del cor"**

- 1.1. Amor romàntic
- 1.2. Matrimoni estable
- 1.3. Idealisme
- 1.4. Afecte i seguretat emocional
- 1.5. Patriotisme
- 1.6. Aventura
- 1.7. Justícia
- 1.8. Independència

## 2. Objectius racionals o "del cap"

- 2.1. Capacitat de resoldre problemes concrets
- 2.2. Promoció o progrés personal
- 2.3. Diners i béns materials
- 2.4. Seguretat econòmica i social
- 2.5. Poder i dominació

El treball és de 1946! Vegeu L. Bardin (1986). *Análisis de contenido* (p. 95). Madrid: Akal.

Finalment, atengueu a aquesta classificació més recent, feta amb òptica psicològica, que treballa sobre el llenguatge descriptiu que empren els nens/es en fer un retrat verbal de les persones. El material de partida van ser narracions etogèniques sobre la base de consignes i en situació d'observació participant.

### 1. Descripció

- 1.1. Indiferenciada (la persona no se separa del seu entorn)
- 1.2. Diferenciada simple (la persona se separa del seu entorn amb criteris simples i globals)
- 1.3. Diferenciada (atribució evident de trets per a la separació)
- 1.4. Disposicional (atribució de trets individuals o de personalitat per a la diferenciació)

### 2. Implicació personal

- 2.1. Egocèntrica (l'altra persona es descriu en funció de característiques pròpies, del narrador)
- 2.2. Mútua (l'altra persona es descriu en termes de la seva relació amb el narrador)
- 2.3. Al·loorientada (no hi ha implicació personal)

Observeu que la categoria 1 té un caire més semàntic ja que està centrada en característiques objectives i prominents de l'entorn. En canvi, les subcategories 2.1 i 2.2 exhibeixen una orientació més pragmàtica en dependre la definició del propi narrador. Pel que fa a la 2.3 torna a ser semàntica en desvincular-se d'aquest mateix narrador.

El sistema no s'ha reproduït complet. Podeu examinar-lo sencer a B. H. Peevers i P. F. Secord (1973). Developmental changes in attribution of descriptive concepts to persons. *Journal of Personality and Social Psychology*, 27, (1), 434-448.



**Les categories AC poden existir ja o ser construïdes en el si d'una investigació, i admeten orientacions sintàctiques, semàntiques i pragmàtiques.**



## Construcció, estructura i funció d'un sistema de categories AC



**Un sistema de categories AC, que classifica material textual i significatiu, pot adoptar una estructura en arbre o combinatòria.**



### ■ Construcció

Abans d'anar més endavant, reitemerem que el procés de construcció d'un sistema de categories AC es pot posar en paral·lel amb el d'un sistema de categories conductual. Ja hem fet notar, però, que hi ha diferències entre ambdós processos que no es poden menystenir perquè deriven de la naturalesa, també diferent, de la matèria primera sobre la qual treballen uns i altres sistemes: textos i registres narratius.

De fet, el material sobre el qual es construeix un sistema de categories AC no és el text en brut, sinó, com també hem vist, el registre d'unitats de text-en-context, pre-classificades, cada una amb el seu codi que ja l'adscriu a un tipus o

classe. És a aquest conjunt de codificacions al qual se solen aplicar els criteris classificatoris.



Resumim a continuació el treball de M. Grawitz en relació amb la històrica campanya de la premsa francesa, amb vistes al referèndum de 1961 sobre la independència d'Algèria.

Lògicament, l'objectiu del referèndum girava entorn de la proporció de respostes afirmatives/negatives. Moltes de les argumentacions dels articles i editorials als diaris es decantaven inequívocament per una d'aquestes dues possibilitats dicotòmiques, de manera que l'eix vertebrador i unificador del futur sistema de categories havia de ser, sens dubte, aquest: a favor del "no" o a favor del "sí".

No obstant això, després d'examinar en detall els textos, es feia palès que hi havia matisos i aparents contradiccions: el "no" podia ser aduït tant per un autor d'extrema esquerra com per un d'extrema dreta o de dreta tradicional, mentre que el "sí" podia provenir de gent d'esquerra moderada, de dreta liberal, etc. A més a més, les argumentacions corresponents mostraven una assertivitat variable, amb més o menys entusiasme en la defensa del vot. Per últim, alguns periodistes i polítics no prenién posició, o recomanaven l'abstenció, el vot en blanc o l'ajornament de qualsevol decisió relativa al tema de fons.

Vist això, un sistema de dues categories "sí/no" hauria resultat força rudimentari. Es va optar per una classificació que recollia els matisos i gradacions de la resposta abans al·ludits:

### 1. Resposta afirmativa.

- 1.1. Entusiasta.
- 1.2. Resignada

### 2. Resposta negativa.

- 2.1. D'esquerra
- 2.2. De dreta

### 3. Abstenció

### 4. Llibertat de vot

### 5. Ajornament

### 6. No pren posició

Així doncs, la classificació solucionava els problemes de diversitat de resposta segregant dos nivells en les dues primeres categories i afegint altres ítems a la dicotomia "sí/no". Les categories generals 1 i 2 representaven les dues direccions generals de l'argument (afirmatiu/negatiu), mentre que les subcategories 1.1, 1.2, 2.1 i 2.2 expressaven la intensitat o la posició dintre de cada una d'aquestes direccions.

Vegeu M. Grawitz (1975). *Métodos y técnicas de las ciencias sociales*, (2), 167-168. Barcelona: Hispano-Europea.

## ■ Estructura

- L'estructura d'un sistema de categories AC ha de satisfer les exigències d'una classificació lògica tal com les satisfien els sistemes de categories conductuals. Recordem tanmateix que aquests sistemes havien d'obeir dos tipus de regles, unes semàntiques o conceptuals i unes altres de temporals i vinculades a la continuïtat del registre. En tots dos casos, però, el sistema havia de complir condicions d'exclusivitat i exhaustivitat que s'expressaven, això sí, de manera diferent en cada modalitat de regla.
- Doncs bé, **en l'elaboració d'un sistema AC les regles temporals d'exclusivitat i exhaustivitat no hi tenen cap paper** com a condicions de classificació.
  - Les categories d'un sistema conductual s'anoten **al llarg del temps**, dins d'una sessió i complint un pla de registre, per la qual cosa poden i deuen obeir criteris d'organització temporal, com el que dues categories no es puguin donar alhora o que no hi hagi intervals de registre buits.
  - ← En canvi, quan la base de la categorització és un seguit d'unitats text-en context, la classificació ha d'ordenar un material fins a cert punt **estàtic**, que no canvia; és la lectura la que introdueix el transcurs temporal, no el text en si.



En el cas de les transcripcions aquesta distinció continua essent certa. Encara que la parla entre dos conversadors flueixi en el temps i hagi estat registrada progressivament al llarg d'una sessió, encara que les seves intervencions s'encavalquin, **la transcripció apareixerà ordenada seqüencialment i estabilitzada com a text**. Això sí: podrem utilitzar una codificació textual que ens indiqui quan hi ha interrupcions o encavalcaments, tal com es fa en l'anàlisi conversacional. *A fortiori*, el mateix podríem dir a propòsit de la transcripció de la parla d'un sol subjecte narrador.

- En conseqüència, **les categoritzacions AC només es regeixen per regles semàntiques o conceptuals d'exclusivitat i exhaustivitat**. Malgrat tot, també la primera apareix sota una llum diferent que en els sistemes conductuals, ja que ara el seu compliment no és exigible en tots els nivells d'organització del sistema, sinó només en el pròpiament categorial. Això fa que **els sistemes AC no sempre adoptin íntegrament la forma de diagrama en arbre, sinó també, almenys parcialment, la de trama combinatòria**. Expliquem aquesta peculiaritat a continuació.

→ **En un diagrama en arbre** com els dels sistemes no AC o conductuals un element d'un nivell superior (objecte d'estudi, categoria, subcategoria, indicador) pot connectar-se amb un o més elements del nivell immediatament inferior (categoria, subcategoria, indicador, unitat de registre); però l'afirmació contrària és falsa: **un element de nivell inferior només pot vincular-se a un element del nivell immediatament superior**; en concret, si és una subcategoria, un indicador o una unitat de registre, només s'enllaçarà a una de les categories que es troben immediatament per sobre.

Aquesta pertinença única és la que garanteix l'exclusivitat en tots els nivells de l'estructura i evita que la classificació sigui inadequadament borrosa de cara al registre o la mesura.

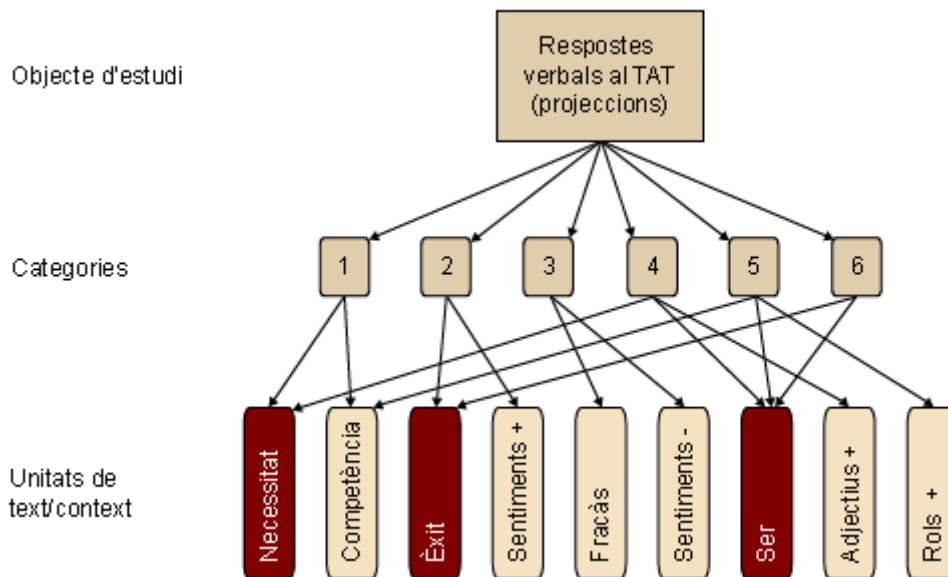
Aquesta mena de relacions són les que hem representat en les figures que hi ha en el subapartat "Estructura d'un sistema de categories i, per descomptat, la que es posa de manifest en l'exemple de Grawitz que ha il·lustrat el punt previ.

- ← Al contrari, **els sistemes AC poden presentar una estructura de tipus combinatòria en bastants casos**, com de fet ja hem vist de passada al principi de l'exposició d'aquest tema en esmentar la Q-anàlisi. Això significa que **un element de nivell inferior pot estar connectat a més d'un element del nivell immediatament superior**; una unitat de registre, indicador o subcategoria es pot enllaçar amb dos o més categories per sobre.

Aquesta llicència, per tant, impedeix el compliment de la condició semàntica d'exclusivitat, ja que un mateix element pot ser contingut per més d'una classe de classificació. Però només ho impedeix en alguns nivells de la classificació i el nivell categorial queda sempre preservat. És a dir: la identitat de cada categoria és exclusiva però, en comptes d'assolir-se mitjançant l'exclusivitat dels elements que conté, aflora gràcies a l'exclusivitat d'una combinació determinada i singular d'aquests elements.

En la figura següent teniu esbossada aquesta relació multívoca **que no fa més que calcar l'organització jeràrquica del llenguatge**. En efecte, el sentit d'una frase no descansa en la possessió exclusiva d'uns mots, ni el d'un mot en una col·lecció única i singular de fonemes o síl·labes. En un i altre cas el que permet reconèixer la forma conspícua i atorgar-li un significat és una combinació particular de components.





Estructura combinatòria de la categorització de les respostes a les imatges del TAT (Test d'Apercepció Temàtica). El primer nivell, de l'objecte d'estudi a les categories, compleix la condició d'exclusivitat. El segon no la compleix perquè les unitats marcades amb color més fosc es connecten amb més d'una categoria d'ordre superior. Compareu la trama de fletxes amb la de la primera figura de l'apartat "Classificació del comportament en l'observació no participant: sistemes de categories interpretatius" del mòdul "Altres tècniques de registre d'elaboració d'informació en observació participant i no participant" d'aquesta assignatura. El sistema no ha estat representat en la totalitat.



Tracteu de dibuixar un esquema semblant al de la figura que teniu en aquest apartat a partir de l'exemple que us proposem. Aquest exemple no prové del camp de l'escriptura, sinó de la comunicació no verbal, però aquesta darrera de vegades regula els signes de manera similar al llenguatge, és a dir, mitjançant regles combinatòries.

Considerem el següent repertori simplificat de categories d'expressió facial centrades en alegria, tristesa i sorpresa, a banda de la neutra o no expressiva.

- **Alegria:** celles alçades en l'extrem exterior o distal, comissures dels llavis alçades (celles+, llavis+).



- **Tristesa:** celles abaixades en l'extrem exterior o distal, comissures dels llavis abaixades (celles+, llavis-).



- **Ràbia:** celles alçades en l'extrem exterior, comissures dels llavis rectes (celles+, llavis0).



- **Neutra :** celles rectes, comissures dels llavis rectes (celles0, llavis0)



Assenyalau, un cop dibuixat l'esquema jeràrquic del sistema de categories, quines parts satisfan el criteri d'exclusivitat semàntica i quines, no complint-lo, tenen una relació combinatòria.



Bàsicament el que heu de fer és traslladar les quatre definicions anteriors de les quatre categories, fetes en quatre

línies diferents segons les normes d'un text convencional, a un esquema o diagrama combinatori com el de la figura que teniu en aquest apartat. En aquest cas (i això ja és dir-vos molt) les regles combinatories només s'apliquen als indicadors, igual que en la figura citada. Per tant, es tracta, com qui diu, de copiar-la, substituint simplement les etiquetes de cada nivell, i tenint en compte —és clar— que el nombre de categories i indicadors no té perquè ser el mateix.

---

## ■ Funcions d'inferència

- La construcció d'un sistema de categories AC implica inevitablement una sèrie de decisions prèvies quant al tipus de relacions que enllacen els diferents nivells d'organització (en l'exemple anterior, les relacions que enllacen "cella alçada" amb "alegria" o "ràbia"). Aquestes relacions poden agafar-se de la teoria, en el cas que ja ofereixin garantia; o poden sortir del sentit comú o del coneixement vulgar; o poden ser pures hipòtesis de treball. En qualsevol cas el que sempre fem és **assignar un significat a la unitat de registre** (assignar un sentit a "cella alçada", el d'alegria o el de ràbia).
- Per consegüent, les relacions entre els nivells per sota de les categories que mesurarem i aquestes categories són de caire semàntic, en tant que representen atribucions de significat. A una unitat de registre se li atorga un significat que la classifica dintre d'una certa categoria. Per això, en determinats àmbits, aquestes unitats amb un vector de significat adreçat a la categoria s'anomenen **indicadors**. Els llavis caiguts són, o poden ser, un indicador de tristesa o depressió.
- És obvi que un dels objectius metodològics de l'investigador ha de ser el de validar aquestes relacions entre indicadors i categories, abans de trobar regularitats o confirmar hipòtesis. L'investigador qualitatiu té al davant dos camins de validació:
  - Un, típicament qualitatiu, és el de consensuar amb l'autor del text i/o amb altres analistes la categorització definitiva. Això no sempre és possible, atès que l'autor pot ser absent o anònim.
  - Un altre és el de recórrer a tècniques multivariants similars a les que serveixen per a dotar els tests de validesa discriminant i consistència estructural (anàlisi factorial, components principals, escalament multidimensional, agrupacions o clusters, alfa de Cronbach, etc.).

Sobra afegir que ambdues vies són, en principi, compatibles.

- Un cop garantida l'estructura del sistema de categories cal plantejar-se una qüestió ben diferent: **quines inferències es poden fer a partir de les mesures realitzades**. Com veurem aviat aquestes mesures (per exemple, la freqüència d'ocurrència) es poden aplicar als elements següents:
  - **Unitats de registre**. Tenen la funció de justificar el sistema en tant que instrument consistent i coherent d'anàlisi, en línia amb el que acabem de comentar als paràgrafs anteriors.
  - **Categories** o variables categorials. Tenen la funció de propiciar explicacions i inferències en comparar parts d'un text o diversos textos.



Us oferim alguns exemples de línies d'inferència, precedides o no per hipòtesis, compatibles amb l'anàlisi de contingut.

- A partir de retalls de premsa, en fase de conflicte prebèlic, deduir i predir si un país A atacarà un país B.
- Sobre la base de la transcripció d'una entrevista saber si el subjecte ha intentat seduir, agradar a l'investigador o, al contrari, li ha fet el boicot i ha mantingut certa distància.
- Establir els trets psicològics de l'autor d'un diari (per exemple, el de Sylvia Plath, poetessa anglesa que es va suïcidar el 1963).
- Detectar el nivell de sinceritat en un autoinforme.
- Reconstruir les línies mestres d'una estratègia de negociació mitjançant l'epistolari entre dos polítics, empresaris, etc.
- Esbrinar en quins aspectes de l'expressió textual es manifesta el racisme o el sexisme en articles periodístics, novel·les, llibres de text, etc.
- Detectar la percepció que té un nen o una nena dels immigrants a partir dels relats escrits que se li han demanat a l'escola.



**Els sistemes de categories AC es construeixen sobre relacions (entre indicadors i categories) i permeten estudiar relacions (entre categories, entre textos, entre autors).**

## La mesura en l'anàlisi de contingut



**La mesura bàsica de l'anàlisi de contingut és la freqüència i es pot aplicar a unitats de registre o codis i a categories.**



### ■ Mesures bàsiques de freqüència i regles d'enumeració

- Acabàvem l'apartat "Construcció, estructura i funció d'un sistema de categories AC" dient que l'anàlisi de contingut mesura en dos nivells ben definits: el de les unitats de registre i el de les categories. I en cada un d'aquests nivells la funció de la mesura és diferent: en el primer serveix per a ponderar la importància o incidència de cada molècula textual, per a legitimar la seva incorporació en una o altra categoria; en el segon, recolza les inferències o permet confirmar hipòtesis de treball.
- Les autèntiques variables són, doncs, les categories, ja que és a partir de les categories que es generen els resultats de la recerca. Són, això sí, variables qualitatives, dicotòmiques o politòmiques, la qual cosa equival a dir que no prenen valors continus, sinó binaris de presència/absència, sí/no, 1/0. Són les variables que apareixien en els exemples que hem anat veient darrerament. Un autor, locutor, emissor, etc., "enuncia fets" o "no enuncia fets"; "fa comentaris" o "no fa comentaris"; "traça una imatge ridícula del pare" o "no traça una imatge ridícula del pare".
- En el si d'aquesta poderosa restricció la mesura més adient i més rendible és la **freqüència absoluta d'ocurrència** o **freqüència** a seques, un recompte dintre de la sèrie de números naturals. Comptabilitza el nombre de vegades que una categoria està present en un text o en zones seleccionades del text.



Adoneu-vos que l'ús conjunt de

1. variables categorials bastides sobre la dicotomia presència/absència i
2. mesures freqüencials d'aquestes variables

és el nucli d'una sintonia harmoniosa entre l'estil qualitatiu i el quantitatiu. Efectivament, ja hem sostingut en un altre lloc que l'ús de taules de presència/absència, per un costat, i de mesures contínues, per l'altre, marquen el contrast principal entre els mètodes qualitius i els quantitius, per la qual cosa la seva utilització conjunta s'ha d'interpretar com un avenç cap a la integració metodològica.

Tanmateix, aquesta promesa d'integració segueix sense eliminar la limitació essencial dels mètodes quantitius i, en particular, de la quantificació freqüencial i altres mesures. I és que no són capaços d'assignar valor a la categoria o el fenomen que no està present, que "brilla per la seva absència". L'enfocament quantitatiu s'enfronta a enormes dificultats per donar sentit al zero.

- Per altra banda, la preeminència de la freqüència també és un tret de la metodologia observacional en general, que també fa descansar les seves mesures sobre categories. Així doncs, com a metodologia de tipus observacional i, dintre d'aquesta, com a metodologia d'observació indirecta o documental, l'anàlisi de contingut és caracteritzat per la utilització, no solament de la freqüència com a mesura principal, sinó d'un lot de mesures derivades o edificades sobre la freqüència, de les quals la **freqüència relativa** o **proporció d'ocurrències** n'és la capdavantera.

Per comoditat, a partir d'ara expressarem la freqüència d'una categoria de contingut C com  $f_C$  i la seva freqüència relativa (que permet estimar probabilitats d'ocurrència) com  $p_C$ :

$f_C$  = nombre de vegades que una unitat de registre codificada o una categoria està present en un text o secció de text.

$p_C$  = nombre de vegades que una unitat de registre codificada o una categoria està present en un text o secció de text respecte a la freqüència total d'ocurrències.

Per tant,  $\hat{p}_C = \frac{f_C}{f_T}$ , essent el denominador el total d'ocurrències. El circumflex sobre la  $p$  indica que es pot tractar d'una estimació de probabilitat.

Si una anàlisi de contingut es fa sobre un sistema de categories S com  $S = \{W, X, Y, Z\}$ , llavors la freqüència total d'ocurrències  $f_T$  serà igual a la suma de les freqüències parcials de cada categoria, de manera que  $f_T = f_W + f_X + f_Y + f_Z$ . Així, si hem obtingut  $f_W = 56$ ;  $f_X = 14$ ;  $f_Y = 23$  i  $f_Z = 102$ , obtindrem  $f_T = 195$  i

$$p_W = 56 / 195 = 0,29$$

$$p_X = 14 / 195 = 0,07$$

$$p_Y = 23 / 195 = 0,12$$

$$p_Z = 102 / 195 = 0,52$$

Naturalment,  $p_W + p_X + p_Y + p_Z = 1$ , com podeu comprovar.

Òbviament, si multipliquem  $p_C$  per 100 obtindrem el percentatge d'ocurrència corresponent a cada categoria (en aquest exemple, 29%, 7%, 12% i 52% respectivament).

- En canvi, per raons que hem adduït fa poc, l'anàlisi de contingut no fa ús de mesures de temps, ja que el text és estàtic des del punt de vista microtemporal. És a dir, malgrat que alguns textos es poden transformar al llarg de mesos, anys o segles (d'esborranys a textos definitius, d'una versió a una altra, d'una formulació de lleis o estatuts a una altra, d'un cos doctrinal a un altre), es donen a l'analista d'un sol cop, no progressivament com es dona el comportament al llarg d'una sessió.

Per consegüent, tampoc no té sentit dir que una categoria AC ha aparegut en un text quinze vegades per minut o que ha ocupat un 24% del temps de registre, afirmacions que sí tenen cap i peus en un registre observacional de sessions: un nen pot canviar d'activitat dues vegades per minut, de mitjana; un obrer pot soldar vint-i-sis peces per hora.

- Ara bé, **la freqüència és una classe de mesura que no té valor si no s'emmarca en un espai o en un temps de referència**. No vol dir gairebé res dir que un nen ha canviat d'activitat dues vegades. Ens preguntarem de seguida en quant de temps ha fet aquests dos canvis d'activitat, perquè no té el mateix significat que els hagi fet en un minut, en una hora o en un dia. Tampoc no té gaire valor l'afirmació "A l'espectacle van assistir 500 persones", atès que aquestes 500 persones poden atapeir un local justament amb aquesta capacitat, o bé omplir tan sols un raconet d'un estadi amb 20.000 seients.

Advertiu, doncs, que és capital **contextualitzar** la mesura, posar-la en un context. Ara, però, aquest context **no és un context d'interpretació, sinó de referència de la mesura**; o, simplement, **un context de mesura**. En el vocabulari de l'anàlisi de contingut diem que no solament cal establir les mesures d'un text sinó també les seves **regles d'enumeració**. Aquestes regles es refereixen al fons o marc en el si del qual la mesura pren significat.

En l'anàlisi de contingut, on ja hem insistit que el temps no serveix de referència, les regles d'enumeració més corrents situen la mesura en relació amb **la línia i la pàgina**. De vegades, les mesures també se situen respecte de **l'oració, el paràgraf, el capítol**.

Podem dir així que una categoria C ha aparegut dues vegades en una línia o set vegades en una pàgina, amb la qual cosa, més que amb una freqüència, treballem amb una **taxa** (tanmateix continuarem mantenint el mateix símbol):

$$f_C = 2 / \text{línia} \text{ ó } f_C = 7 / \text{pàgina}.$$

Igualment podem afirmar que una categoria C ha aparegut una mitjana de 1,02 vegades per línia o de 5,83 vegades per pàgina:

$$\bar{f}_{C/LÍNIA} = 1,02 ; \bar{f}_{C/PÀGINA} = 5,83$$

Podríem haver donat aquestes freqüències i mitjanes per oració, paràgraf o capítol però, tot i que aquests denominadors poden estar justificats en certes condicions, no permeten una estandardització de la mesura tal com la fan possible els marcs de referència espacial, fixos, com la línia o la pàgina. Si considerem la freqüència respecte dels contextos d'extensió variable com l'oració, la taxa resultant és difícil d'interpretar.

Com és natural podem calcular també si ens interessa desviacions típiques, variàncies, etc.

## ■ Altres mesures derivades de la freqüència o que la pressuposen

- **Freqüència de contingència o co-ocurrència**. És una variant de la freqüència d'ocurrència per la qual sabem el nombre de vegades que dos o més unitats de text coincideixen en un mateix context, o el nombre de vegades que una unitat de registre codificada o una categoria coincideix en un mateix context de mesura.

Per tant, podem utilitzar aquesta mesura per fer afirmacions com que la unitat textual "drets" i la unitat textual "igualtat" han coincidit vint-i-un vegades en el context d'oracions que són enunciats de lleis. Formalment, si "drets" l'anomenem w i "igualtat" l'anomenem x, i aquest tipus d'oració el designem per L, tindrem el següent:

$$f_{w,x/L} = 21.$$

O, igualment, podríem servir-nos d'aquesta freqüència per a dir que la categoria Y i la categoria Z han coincidit disset vegades en una línia o trenta-quatre vegades en una pàgina, ara referint-nos al context de mesura més que al d'interpretació.

- **Freqüència ponderada**. L'ús de la freqüència com a mesura treu partit d'una sèrie de supòsits, dels quals un dels més importants és el següent: en principi totes les unitats o categories són equivalents i virtualment tenen la mateixa probabilitat d'aparèixer al text. És com si diguéssim que totes les categories arrenquen de la mateixa línia de partida en una cursa i cap d'aquestes categories s'avança a l'altra fins que no comença la cursa. Sobre la base d'aquest supòsit inicial la rellevància d'una unitat o categoria creixerà a mesura que en creixi la freqüència.

Aquest supòsit, però, topa sovint amb diverses limitacions. Ja n'hem citat una: de vegades la circumstància que un element del text tingui freqüències baixes o no aparegui és tan significatiu com el seu contrari.

Un altre obstacle, que comentarem a continuació és que aquesta condició d'equivalència pot, senzillament, no complir-se, atès que hi ha mots, frases, expressions, temes que, en un determinat tipus de text, tenen més presència, *a priori*, que altres. Àdhuc, podem fer una estimació teòrica de la incidència d'aquests elements en el text ja que hi ha, sortosament, registres d'ús en moltes llengües que ens diuen quina és la freqüència amb què podem esperar trobar-los. Aleshores la incidència d'un mot o d'una categoria en un text s'ha de posar en relació amb la seva incidència estadística mitjana en materials estàndard del mateix tipus.

A causa de raons com la que acabem d'exposar, i a d'altres semblants, sovint s'utilitzen mesures de **freqüència ponderada** o, el que és el mateix, **mesures de freqüència en què cada unitat o categoria registrades no té el mateix "pes" o rellevància, sinó que unes en tenen més que d'altres**. El pes és assignat multiplicant cada freqüència empírica per un valor decidit prèviament, a partir de consideracions teòriques o de registres anteriors com els que indicàvem més amunt.

Si tenim un sistema de categories  $S = [X, Y, Z]$  i considerem que Y té el doble de pes que X, i Z el triple, aleshores les freqüències ponderades  $f_{\text{pond}}$  que comptabilitzarem seran les següents:

$$f_{\text{pond } X} = f_X ; f_{\text{pond } Y} = 2(f_Y) ; f_{\text{pond } Z} = 3(f_Z)$$

si  $f_X = 61$ ,  $f_Y = 20$ ,  $f_Z = 14$ , aleshores

$$f_{\text{pond } X} = 61 ; f_{\text{pond } Y} = 2 \times 20 ; f_{\text{pond } Z} = 3 \times 14.$$

- **Intensitat.** Un desafiament seriós dins de l'anàlisi de contingut és la mesura de la intensitat, essencialment vinculada a les categories semàntiques. Aquest desafiament, ara per ara, és difícilment afrontable fora de procediments psicomètrics.

Assumir la mesura de la intensitat vol dir, en el fons, transcendir el plantejament dicotòmic o binari de la quantificació categorial i passar, com a mínim, a una quantificació politòmica. O sigui, que les coses, si es desitja mesurar la intensitat d'una categoria, no són blanques o negres, sinó que poden oferir una gamma de grisos, gamma graduada en forma d'escala ordinal si més no.

A tall d'il·lustració compareu aquests tres enunciats corresponents a tres unitats de registre classificables en una categoria que podríem etiquetar com "rebuig de la política presidencial" per part de polítics de l'oposició:

- "Podríem arribar a considerar la possibilitat de desaprovar la política del president".
- "Cal denunciar, amb pena, l'actual orientació política del president".
- "Certament, en el passat hem estat sovint en desacord amb la política del president".

És clar que convindria situar cada una d'aquestes afirmacions en una escala d'intensitat quant al grau de rebuig que expressa o vol expressar. Un procediment artesanal (però força emprat) seria assignar ponderacions a cada ítem a), b) i c), tractant-los com si fossin subcategories ordenades d'una categoria general de rebuig. Els criteris on ens recolzaríem podrien ser diversos i combinables: semàntica dels verbs tal com els defineix el diccionari o els usuaris, temps d'aquests mateixos verbs, adverbis –especialment els de mode– que reforcen o suavitzen el contingut verbal, adjectius qualificatius, signes d'admiració. En l'exemple anterior podríem assajar la ponderació  $a = 2$ ,  $b = 3$ ,  $c = 1$ .

Més enllà de procediments com aquest, la construcció d'una escala d'intensitat obliga a aplicar totes les fases i requisits d'un escalament psicomètric.

- **Direcció.** Bona part del que hem dit a propòsit de la intensitat és també aplicable a la direcció.

Moltes interpretacions del sentit o valor d'un text es tornen més riques i completes si se situen en una dimensió de mesura, entre un mínim i un màxim. Per exemple, en un estudi sobre integració de nens i nenes immigrants en la població autòctona podríem servir-nos d'una categoria de "Desig de convivència" i mesurar-la entre un desig mínim, zero o negatiu, i un de màxim, amb un nombre de valors intermedis.

Històricament, aquestes mesures s'han resolt a partir d'escals d'estimació, qüestionaris que el subjecte ha de respondre escollint un valor dins d'una sèrie, d'un a cinc, d'un a set, d'un a deu, etc. C. E. Osgood ha estat un



dels paladins d'aquest tipus d'estratègia.

Una altra via de resolució de les mesures de direcció és la creació d'índexs amb un àmbit de variabilitat adequat a aquest tipus de recerca. De seguida en veurem un exemple.

- **Ordre seqüencial.** L'ordre amb què es manifesten seqüencialment unitats de registre o categories és un altre vessant de la informació que admet quantificació. En aquest terreny l'anàlisi de contingut comparteix recursos, un cop més, amb la metodologia observacional, en el si de la qual s'ha desenvolupat una munió de tècniques destinades a la mesura i modelització de l'ordre dels elements d'una sèrie o seqüència.

En un enfocament d'anàlisi de contingut ens pot interessar saber el grau d'organització seqüencial d'un text, les regles subjacents, el patró nuclear que li confereix la seva forma aparent. En perspectiva plenament qualitativa l'investigador pretindrà sovint descobrir la matriu de ritual o cerimònia que conforma el material verbal o textual, tal com es posa especialment de relleu en els enfocaments etogènics i etnometodològics.

L'estratègia essencial de la cerca de patrons d'ordre consisteix en el còmput de **freqüències de transició**. Aquestes són les freqüències amb què cadascuna de les categories d'un sistema AC segueix en la seqüència textual a la resta de categories del sistema, inclosa ella mateixa. En general, doncs, són freqüències de parells d'elements contigus,  $f_{IJ}$ . De fet, es podrien computar tant freqüències de precedència (I precedeix J) com de seguiment (J segueix I), o ambdues a la vegada.

Des del punt de vista qualitatiu, la categoria que precedeix una altra categoria és el context antecedent, i la que la segueix, el context conseqüent. En les línies prèvies hem estat sempre al·ludint a precedències o seguiments immediats, però això no té perquè ser així: les dependències entre els elements d'una cadena textual es poden establir a més distància, entre una categoria i la que la precedia o la que la seguia dos, tres o quatre posicions més enrere o més endavant.

Aquestes freqüències de transició poden ser transformades en probabilitats condicionades, que són les que pròpiament ens permeten establir les regularitats de la seqüència.



Per aquesta via poden arribar a saber, posem per cas, el següent:

- Que en una entrevista a una mare cada pregunta sobre els seus fills tendeix a anar seguida, en primera o segona opció, de categories de resposta classificables com a "afectives" i, després, en tercera posició, per una negació de les anteriors,
- Que en un diari adolescent els inicis de les anotacions solen ser expressius quan el text continua després narrant conflictes o drames (familiars, escolars, sentimentals), mentre que s'inclinen a la descripció quan els esdeveniments que omplen la pàgina d'aquell dia són més aviat neutres.

- **Índexs.** Un índex molt popular és el de **diversitat o riquesa de les unitats o categories (Id)**. Tradicionalment s'ha emprat per a calcular-lo una senzilla **proporció entre el nombre d'unitats o categories diferents (D) i el nombre total d'unitats o categories (T)**:

$$Id = \frac{D}{T}$$

Si invertim el quocient, tindrem un **índex de pobresa o monotonia, Im**:

$$Im = \frac{T}{D}$$



Considerem la frase que acabeu de llegir, és a dir: "Un índex molt popular és el de **diversitat** d'unitats o categories. Tradicionalment s'ha emprat una proporció senzilla entre el nombre d'unitats o categories diferents i el nombre total d'unitats o categories" (l'hem escurçat una mica i hem suprimit els símbols inicials Id, D i T).

En aquest segment el valor d'Id per a unitats textuais tipus "mot" seria  $Id = 23 / 35 = 0,657$  ("s" i "d" han estat comptats com un mot). El valor d'Im seria  $Im = 35 / 23 = 1,52$ .



Actualment aquesta classe d'índexs es calcula traient partit de la teoria de la informació. En l'àmbit de l'observació de camp del comportament sabem que hi ha una relació entre D i T, expressada en l'equació  $D = CT^z$ , on C i z són les constants que, precisament, cal esbrinar en cada investigació. Aquesta relació es pot transformar en lineal sobre escala logarítmica, ja que

$$\log D = \log C + z \log T$$

Per tant, en una gràfica logarítmica la relació entre el nombre de categories o unitats diferents i el nombre total de registrades fins a un cert moment és una recta, que es pot dibuixar mitjançant tècniques de regressió simple.

Una altra família d'índexs popular és la dels dedicats a quantificar la relació entre la freqüència d'ús de diferents categories gramaticals. Vegeu-ne dos exemples: el quocient **adjectius/verbs** (amb denominador habitualment fixat en 100) ha permès distingir la locució esquizofrènica de la normal o no patològica. Per a alguns autors el quocient **(noms + verbs)/(adjectius + adverbis)** és un bon indicador del tarannà del discurs polític, dinàmic o descriptiu segons que el trencat superi la unitat o en quedi per sota.



**Les freqüències i altres mesures derivades s'han de contextualitzar mitjançant regles d'enumeració.**

## Exemples, programes, síntesi

- Exemples abreujats d'anàlisi de contingut
- Tractament informatitzat de dades textuais i de contingut
- Programes d'ordinador per a l'anàlisi de contingut
- Diagrama d'operacions bàsiques en l'anàlisi de contingut

## Exemples abreujats d'anàlisi de contingut



**Una anàlisi de contingut completa ha de cobrir les fases de mostratge, creació d'unitats, categorització i mesura, abans de passar a les inferències.**



### ■ Exemple 1

- Escollim per a aquesta primera exemplificació un dels relats etogènics que hem utilitzat per tal d'il·lustrar aquesta orientació documental. Recordeu que es tractava d'un paràgraf en què un marit propietari d'un pis expressava estranyesa respecte a l'incompliment de les expectatives de venda de l'immoble, expectatives que, tant ell com la seva dona, s'havien fet després de la visita d'una parella de pretesos compradors.



Podeu veure l'exemple del subapartat "Els relats d'episodis socials. La perspectiva etogènica" del mòdul "L'observació documental" d'aquesta assignatura.



Per exemple, els gestos adreçats a objectes es podrien vincular a estats agressius abocats a l'exterior del subjecte, mentre que els gestos de contacte corporal amb el propi cos podrien ser l'expressió d'una violència intrapunitiva o dirigida contra el mateix cos. El contingut dels deliris esquizofrènics varia segons correspongui a una fase aguda o crònica i segons el grau d'integració o autisme.

Sobre aquest material assajarem les principals operacions que l'anàlisi de contingut realitza sobre un text i que fins ara hem anat enumerant.

Incloem de nou el text en qüestió:



"Aquesta gent, quan ens va avisar l'agència de que ens els enviaven, van venir de seguida. Gairebé no ens van donar temps a posar el pis en ordre. Semblaven molt interessats i semblaven bons compradors. Al principi jo hagués jurat que estaven a punt de decidir-se. De fet fins i tot van parlar (ells dos) de canviar algunes coses del pis. Van comentar que les finestres les farien totes noves. Vam pensar que tenien diners llargs i que serien bons pagadors. Van estar molta estona a casa, i quan van marxar semblava que ja s'hi trobaven com a casa seva. Jo vaig dir a la Cíntia (la meva dona) que truqués a l'agència perquè els collés i fossin a l'aguait. Però no van dir res en tot el dia. Ni l'endemà, ni la resta de la setmana. No enteníem res ni l'agència, tampoc. Al final els vaig trucar personalment i a sobre crec que els va molestar. Ja no els interessava. Qui ho entén!"

Adoneu-vos primer de tot que el text presenta dues parts ben diferenciades: una en què el narrador expressa o creu confirmar diverses expectatives de venda, i una altra en què palesa el seu desencís en veure aquestes expectatives frustrades.

- Les **unitats de registre** poden ser **mots** o **frases** en **context d'oracions**. Els mots són indicadors categorials que prenen tot el sentit en el context de les oracions. Les oracions estan delimitades o definides operacionalment, a) quan són simples, pel fet d'estar compreses entre punts i seguit consecutius, o entre l'inici i punt i seguit, o entre punt i seguit i punt final; b) quan són compostes copulatives, per estar compreses entre un punt i el connector "i" (si són compostes d'altres tipus es consideren simples).
- Les categories es basen en diferents mots o frases –els indicadors– i són de dos tipus: **categories d'expectativa confirmada, E<sup>+</sup>**, i **categories d'expectativa frustrada, E<sup>-</sup>**.

Així, les frases "molt interessats" i "bons compradors", en els contextos corresponents de "Semblaven molt interessats" i "semblaven bons compradors" es codifiquen com E<sup>+</sup>. "No van dir res" en el context de "Però no van dir res en tot el dia" es codifica com E<sup>-</sup>. Etc.



Ometem en aquesta exposició simplificada altres subtileses de la codificació, com la influència que de cara a la interpretació categorial pot tenir un marc de referència més ampli, com el de les altres oracions que, a dreta i esquerra, envolten un mot i el context immediat.

Per altra banda, notareu que, en alguns casos, la unitat textual, mot o frase, pràcticament coincideix amb el context.

- Amb tot això arribem a la següent llista de codificacions i, per tant, de categories (indicadors en negreta):
  - Aquesta gent, quan ens va avisar l'agència que ens els enviaven, **van venir de seguida**: E<sup>+</sup>
  - Gairebé **no ens van donar temps** a posar el pis en ordre: E<sup>+</sup>
  - Semblaven **molt interessats**: E<sup>+</sup>
  - semblaven **bons compradors**: E<sup>+</sup>
  - Al principi jo hagués jurat que **estaven a punt de decidir-se**: E<sup>+</sup>
  - De fet fins i tot van parlar (ells dos) de **canviar algunes coses del pis**: E<sup>+</sup>
  - Van comentar que les **finestres les farien totes noves**: E<sup>+</sup>
  - Vam pensar que **tenien diners llargs**: E<sup>+</sup>
  - (vam pensar) que **serien bons pagadors**: E<sup>+</sup>
  - Van estar **molta estona** a casa: E<sup>+</sup>
  - quan van marxar semblava que ja **s'hi trobaven com a casa seva**: E<sup>+</sup>
  - Jo vaig dir a la Cíntia (la meva dona) que truqués a l'agència **perquè els collés**: E<sup>+</sup>
  - (Jo vaig dir a la Cíntia (la meva dona) que truqués a l'agència perquè) **estiguessin a l'aguait**: E<sup>+</sup>
  - Però **no van dir res** en tot el dia: E<sup>-</sup>
  - (...no van dir res en tot el dia) **(Ni) l'endemà**: E<sup>-</sup>
  - (...no van dir res en tot el dia) **(ni) la resta de la setmana**: E<sup>-</sup>
  - **No enteníem res**: E<sup>-</sup>
  - l'agència, **tampoc**: E<sup>-</sup>
  - Al final **els vaig trucar personalment**: E<sup>-</sup>
  - a sobre crec que **els va molestar**: E<sup>-</sup>
  - Ja **no els interessava**: E<sup>-</sup>
  - **Qui ho entén!**: E<sup>-</sup>

- A partir d'aquesta codificació de categories podem emprar diverses mesures:

$$f_{E^+} = 13; f_{E^-} = 9; p_{E^+} = 13 / 22 = 0,59; p_{E^-} = 9 / 22 = 0,41$$

Podem establir direccionalitat mitjançant un índex senzill com

$$I = \frac{f_{E^+} - f_{E^-}}{f_{E^+} + f_{E^-}}; \text{ de manera que } I = \frac{13 - 9}{13 + 9} = 0,18$$

L'índex està comprès entre -1 i 1, es a dir:  $-1 > I > +1$ . Entenem, doncs, que en aquest segment de text, sense considerar l'ordre seqüencial, les expectatives positives dominen lleugerament sobre les negatives.

Quant a l'índex de diversitat  $I_d$ , referit a categories, és fàcil de computar atès que només tenim dues categories i ja sabem el nombre total de categories comptabilitzades,  $f_{E^+} + f_{E^-} = 13 + 9 = 22$ . Així

$$I_d = 2 / 22 = 0,09$$

Pel que fa a l'ordenació seqüencial veiem, a vista d'ocell, que aquest paràgraf presenta una organització molt rígida i dicotomitzada. Totes les categories  $E^+$  s'agrupen en les vuit primeres línies, mentre que les  $E^-$  es concentren en les tres finals. Això es deu a la retòrica de "fer-se la víctima" que utilitza el narrador i que consisteix a insistir molt al començament en la confiança en un curs d'esdeveniments positius (per a ell) per passar després, abruptament, a exposar la desil·lusió terminal. Així, el lector rep al final la mateixa "dutxa d'aigua freda" que ha rebut ell.

Vista aquesta ordenació tan rígida, el sentit comú qualitatiu ens diu que sobra fer cap còmput de freqüències de transició. No obstant això, per tal d'amortitzar l'exemple, vegeu com es computarien aquestes freqüències  $f_{ij}$  mitjançant una taula de contingència seqüencial:

		Categoria consegüent		
		E <sup>+</sup>	E <sup>-</sup>	Total
Categoria precedent	E <sup>+</sup>	12	1	13
	E <sup>-</sup>	0	8	8
Total		12	9	21

La lectura d'aquesta taula ens diu que la **transició o parell seqüencial E<sup>+</sup>E<sup>+</sup>**, s'ha produït 12 vegades, **E<sup>+</sup>E<sup>-</sup>**, 1 vegada, **E<sup>-</sup>E<sup>+</sup>**, 0 vegades i **E<sup>-</sup>E<sup>-</sup>**, 8 vegades. L'única transició **E<sup>+</sup>E<sup>-</sup>** articula les dues parts del text a les que més amunt fèiem al·lusió.

Les distribucions de freqüència en taules de contingència com aquesta es poden contrastar amb diferents models de distribució aleatòria mitjançant proves tipus  $\chi^2$ .



Vegeu V. Quera (1993). Anàlisi secuencial. A *Metodologia observacional en la investigació psicològica*, (2). *Fundamentación* (2), 341-586. Barcelona: PPU

- Noteu que totes les inferències que ens han permès les mesures apunten a l'autor individual del relat.

## ■ Exemple 2

L'exemple que us oferim ara es presentarà amb menor detall però, en compensació, segurament oferirà més versemblança en tractar-se del resum d'una investigació sencera. N'hem extret allò que era inseparable de les decisions metodològiques essencials. El punt de vista serà més general i això ens permetrà distingir amb més claredat les fases típiques d'una recerca d'anàlisi de contingut tal com han estat seqüenciades aquí.

- **Objecte d'estudi i hipòtesis.** L'autor es proposa estudiar un moviment de caire religiós anomenat "Meditació Transcendental" mitjançant els seus documents i –també– registres d'observació participant.



L'article de referència és E. Woodrum (1984). "Mainstreaming" Content Analysis in Social Science: Methodological Advantages, Obstacles and Solutions. *Social Science Research*, 13, pp. 1-19.

La part que resumim ara es troba entre les pàgines 8 i 18.

Més concretament el focus de l'estudi apunta a l'expressió de creences en els textos i a l'exemplificació de rols socials i obligacions assumibles pels feligresos. Se sosté la hipòtesi següent: amb el temps, els continguts dels textos, tant pel que fa a les creences com a les obligacions, anirà derivant cap a **una popularització dels temes**.

"Popularització" vol dir que els continguts religiosos esdevenen menys espirituals i més abocats a les coses terrenals i a la pràctica quotidiana. Aquest matisos es capten mitjançant les referències als beneficis mundans de l'ortodòxia, al rendiment personal i social de rituals i cerimònies, a l'accessibilitat de Déu, a la inutilitat de certs conflictes i esforços i, per contra, a la utilitat de les actituds conservadores...

Els valors de la variable temps, essencials per tal d'avaluar aquests canvis temàtics, estan garantits per l'any de publicació dels llibres, revistes i documents examinats.

- **Selecció i mostratge.** Pel que fa al **mostratge intertextual** els textos se seleccionen dins d'un període de vint anys, des de la fundació del moviment (1961) fins als primers símptomes de crítica.

S'accepten dins del corpus tots aquells materials que es puguin considerar documents **oficials** de la doctrina i que descriguin creences, pràctiques i objectius socials. Més concretament, són oficials els documents que compleixen almenys una d'aquestes condicions: a) ser publicats per la premsa privada del moviment; b) ser escrits per un representant conegut i acceptat del moviment; c) ser recomanat en una de les llistes enviades als militants i simpatitzants del moviment des de la seu central.

Sobre aquesta base es recopilen 5716 pàgines de literatura sobre meditació transcendental, però el període d'estudi queda reduït a **setze anys**, interval durant el qual es va publicar material que satisfia almenys una de les tres condicions *a*, *b* o *c*.

El **mostratge intratextual** es resol per procediments aleatoris i quotes proporcionals a l'extensió de cada document del corpus. D'aquesta manera s'escullen **vint-i-cinc paràgrafs per any** i, per tant, el material definitiu subjecte a anàlisi és de vint-i-cinc x setze = **quatre-cents paràgrafs**.

- **Unitats d'anàlisi. La unitat textual és el mot.** L'anàlisi treballa sobre un inventari de **vint-i-sis mots** que, situats en el context serveixen d'indicadors de les categories que de seguida esmentarem.

**La unitat de context (interpretatiu) és l'oració.**

Per tant, **la unitat de registre text-context és el mot-oració.**

- **Categories.** L'autor troba **dinou categories temàtiques** construïdes a partir de diverses **combinacions de les vint-i-sis unitats de registre, un cop codificades**, les quals funcionen com a indicadors semàntics.

L'autor fonamenta el valor d'aquests indicadors en els estudis anteriors sobre la distinció entre religió popular o terrenal, per un costat, i mística o espiritual, per un altre. Per exemple, "il·luminació" i "esperit" són indicadors de categories de religió espiritual, mentre que "feina" o "salut" ho són de categories de religió popular o popularitzada, més adreçada als interessos seculars.

**El criteri d'exhaustivitat no es compleix totalment**, ja que en un 3% del corpus textual (els quatre-cents paràgrafs) no hi apareix cap de les dinou categories previstes.

- **Mesures.** Es fan **mesures de freqüència a dos nivells**, comptant els **indicadors o unitats de registre** (mots-oració) i comptant també les **categories temàtiques**.

Com a regla d'enumeració o **context de referència de la mesura es fa servir el paràgraf**.

Tenim, doncs, mesures de freqüència de mot/paràgraf i de categoria/paràgraf en cada interval anual considerat. D'aquesta manera es pot validar la relació entre unitats de registre i categories alhora que s'estableixen dos nivells d'anàlisi de la popularització de temes religiosos.



No hem inclòs aquí cap comentari sobre els procediments seguits per dotar de fiabilitat i validesa a aquestes mesures. La raó és que les qüestions relatives a la qualitat i consistència de les dades encara no les hem tractat. Ho farem en el mòdul "La qualitat de les dades en la investigació qualitativa". Tanmateix, direm que la consistència interna tant del repertori de mots com del de categories està garantit pels valors obtinguts de l'alpha de Cronbach.

No es calculen mesures d'intensitat o direcció.

- **Resultats i anàlisi.** La hipòtesi de l'autor se sosté gràcies al següent:
  - L'alta correlació entre l'any de publicació i les mesures de les categories temàtiques: les categories "espirituals" minven respecte de les "seculars" a mesura que transcorren els anys de publicació. La correlació entre anys i mesures de mots no és tan significativa.



De fet no es comparen directament les mesures de freqüència sinó índexs calculats separatament per a categories i mots. Aquests índexs són, precisament, les alpha de Cronbach que, com vèiem més amunt, asseguren la consistència interna dels llistats de mots-context i categories.

- L'anàlisi de la variància, amb nivells de confiança superiors també en el cas de les categories temàtiques.
- **Conclusions. Els textos oficials de la meditació transcendental han anat popularitzant els seus continguts al llarg del període d'estudi.**

Noteu que totes les inferències que ens han permès les mesures apunten a una institució com a autora dels textos.



Tornem al text de l'exemple 1, és a dir, al relat del desencís d'un marit que creia que estava a punt de vendre el pis. Intenteu fer un exercici semblant al que hem realitzat, a partir del mateix text, però utilitzant ara un altre sistema de categories, també de dos elements. Aquest sistema us el proposem nosaltres, i és de caire un xic diferent que el que hem emprat en l'exemple.

Les dues categories que us proposem ara són: **enunciats factuais** i **enunciats virtuals**.

**Els enunciats factuais F** expressen, comuniquen o denoten fets observables, que no depenen de l'opinió o la percepció particular del subjecte –en aquest cas, el marit. Poden anar acompanyats de frases com "De fet", o no anar-hi.

**Els enunciats virtuals V** depenen del punt de vista o de la percepció del subjecte, expressen o comuniquen **creences**. Solen anar acompanyats dels verbs "semblar", "creure", "pensar", etc., però poden igualment no anar-hi.

- A partir d'aquests nous criteris feu un nova codificació de les unitats de registre de l'exemple 1.
- Sobre la base del llistat obtingut, **calculeu les mateixes mesures** que s'han tobat en l'exemple de referència.



Pel que fa a la nova codificació, podeu utilitzar la mateixa tàctica de segmentació que hem fet servir en l'exemple comentat, o sigui, la de distingir segments separats per punts i seguit o pel connector "i". Per tant, ho teniu fàcil: com que partireu de les mateixes unitats de registre, l'únic que heu de fer és codificar-les com enunciats factuais o virtuals, enganxar-los l'etiqueta **F** o la **V**.

Pel que fa a les mesures, us donem la freqüència com a punt de partida:  $f_F = 15$ ;  $f_V = 7$ . Ara bé, tenint present que, de vegades, la interpretació d'una categoria pot diferir d'un observador a un altre, si no hi ha un procés d'entrenament i consens, no us desanimeu si no obteniu exactament aquestes freqüències.

En qualsevol cas, seguïu vosaltres amb la resta de mesures preses en l'exemple 1.





Una recerca d'anàlisi de contingut pot donar peu a inferències sobre autors individuals (exemple 1) o sobre institucions (exemple 2).

## Tractament informatitzat de dades textuais i de contingut



L'anàlisi de contingut informatitzat redueix la distància entre l'enfocament qualitatiu i el quantitatiu.



### Generalitats sobre la informatització de l'anàlisi de contingut

- Des dels anys cinquanta l'estudi de textos ha estat ajudat pels ordinadors, però és a partir dels seixanta quan comencen a aparèixer en el mercat programes específics d'anàlisi com el General Inquirer (1966) o el Words (1975).

Tanmateix aquests primers programes són dissenyats per a la pura quantificació. És ja en la dècada dels setanta del segle XX que trobem els primers programes d'anàlisi de contingut amb vocació qualitativa. Des de llavors aquests programes s'han anat desenvolupant en diferents direccions i amb un ritme prou accelerat. Tots ells tenen una versió principal destinada a ordinadors personals mentre que alguns disposen també de versions perquè s'utilitzin de manera compartida en xarxa.

- L'ús de programes d'anàlisi de contingut i, en general, d'anàlisi textual ha tingut, entre altres conseqüències més òbvies, la de fer trivials algunes de les marques diferencials dels mètodes qualitius respecte dels quantitius, i viceversa. Ja hem fet notar que, en qualsevol cas, aquestes diferències no es podien localitzar, grollerament, en la simple decisió de "fer o no fer números". Per començar, quan l'analista veu facilitada la tasca de computar, classificar i tractar estadísticament els registres, l'elecció entre una o altra metodologia no depèn dels tipus de coneixements que té (més o menys matemàtics) ni de l'embarum de feina que està disposat a assumir, sinó de conviccions metodològiques.

Però, a més d'aquesta clarificació del que són –estrictament– les preses de posició metodològiques, és igualment cert que, en la pantalla, la línia divisòria entre ambdues opcions no és tan gruixuda; més encara, en aquesta situació, la utilització de tècniques quantitatives en el si de plantejaments qualitius no sembla cap contradicció. Ja haureu advertit que hem parlat de quantificació i de mesures com a suport de l'anàlisi simplement categorial o qualitatiu.

En resum, l'ús de programes d'anàlisi de contingut indueix una certa convergència metodològica: l'investigador qualitatiu té al seu abast recursos quantitius i orientats a la recerca nomotètica complementaris que, per inèrcia, tendiria a utilitzar; i l'investigador quantitatiu es pot col·locar amb comoditat en la perspectiva de les dades individualitzades i les diferències entre casos, a partir de taules de presència/absència de trets.

### Fases del procés estàndard d'anàlisi de contingut per ordinador

- Tots els programes actuals d'anàlisi textual i de contingut poden fer una sèrie de tasques elementals i indispensables com marcar el text, identificar segments, recuperar-los, computar-ne les freqüències, classificar-los a partir de la codificació i després mesurar-la freqüencialment, relacionar unitats i categories, etc. De fet els processadors de textos d'avui dia també poden fer algunes d'aquestes funcions.

Al principi, els programes que treballaven amb MS-DOS exigien la introducció prèvia del text en un processador per tal de convertir-lo en ASCII. En aquests moments, en canvi, l'ús de Windows o d'Hypercard com a sistemes operatius fa possible l'acceptació directa del text per part del programa.

- Els passos que cobreix habitualment una anàlisi de textos informatitzada són, si fa no fa, els mateixos que donen forma al seu esquelet metodològic:



Seguim P. Navarro, C. Díaz (1995). "Análisis de contenido". A J. M. Delgado, J. Gutiérrez. *Métodos y técnicas cualitativas en ciencias sociales*, pp. 177-224. Madrid: Síntesis.

- Establiment del corpus textual.** El programa ens pot ajudar a engalzar els diferents documents recopilats o

produïts, creant un **text únic** sobre el qual treballar.

- **Marcatge d'unitats de registre i codis.** L'ordinador sovint serveix el material en línies separades i numerades. El marcatge el pot fer l'analista assenyalant el principi i final de la unitat en el text i etiquetant-la amb un símbol. Un segment de text codificat es pot superposar a altres o estar inclòs (imbricat) en un de més llarg; per tant, **la codificació pot ser múltiple**, ja que un segment pot rebre més d'una marca.

De vegades es poden crear rutines que permeten que la màquina marqui tota sola el segment registrable a partir de les instruccions que se li han donat (per exemple, sobre un text constitucional, suposem la instrucció < Etiqueta amb Q totes les ocurrences del mot "drets" que vagin acompanyades 10 mots abans i/o després per l'arrel "igual-" >). Molts programes ja tenen models de codificació previstos que poden ser utilitzats o no per l'usuari.

- **Categories i classificacions.** A partir de l'etapa anterior el programa pot subagrupar les codificacions en diverses llistes, segregant nivells d'organització si cal. Si hi ha una part del text no codificat, el separa. En aquestes agrupacions constarà la posició original que ocupava cada segment textual abans de passar a la llista del codi corresponent, facilitant la recuperació d'informació. També hi constaran d'alguna manera les codificacions múltiples.

Fins aquí els programes ofereixen funcions útils sobretot per a estudis descriptius. Són fonamentalment recuperadors de text (*text retrievers*). Una altra opció ben diferent són les funcions que permeten omplir de contingut semàntic cada codificació, transformant-la en una autèntica categoria conceptual. Llavors el sistema de codificació establert es pot estimar com un sistema de categories o prendre's com la base del sistema. Si es vol treballar amb alguna profunditat teòrica cal utilitzar programes capaços de relacionar conceptes semànticament; és a dir, programes que no sols assenyalen el segment sinó que li assignen un contingut abstracte (*Nudist* en seria un exemple).

Gràcies a la codificació múltiple la majoria de programes poden detectar coocurrències en el si de contextos d'interpretació o de mesura i, a partir d'aquestes coocurrències, establir les relacions corresponents. A més d'això alguns programes poden classificar en diversos nivells les fonts de procedència (institucions, autors, entrevistats) a partir de les variables que se'ls indiquin (edat, sexe, ideologia, classe social, nivell d'estudis, etc.). Sobre aquesta base, llavors, poden buscar relacions de coocurrència entre categories textuales o de resposta i categories de mostra.

- **Mesura i anàlisi.** Ja hem dit que les possibilitats de computació de freqüències o de càlcul d'índexs d'aquests programes sempre són altes (en realitat, és el més fàcil per al programa). A més a més, molts fan possible la creació de matrius de dades i tenen sortides a fulls de càlcul i paquets estadístics tipus SPSS.



**L'anàlisi de contingut ajudada per ordinador inclou els mateixos passos que la que es fa sense ajut informàtic; la diferència rau en la facilitat del tractament del material.**

## Programes d'ordinador per a l'anàlisi de contingut



**Els principals programes d'anàlisi textual permeten realitzar ràpidament i efectivament totes les fases d'una anàlisi de contingut estàndard.**



### ■ Programes principals d'anàlisi qualitativa de contingut

Ens cenyirem a programes d'anàlisi de contingut pròpiament dit i ens situarem en una òptica qualitativa bastant rigorosa. Altrament, aquest apartat podria ser inacabable, atès el nombre de programes de tractament i anàlisi de textos actualment disponibles. La selecció que hem fet procura oferir exemples de diferents tipus de prestacions dintre de l'òptica escollida, des de la transcripció fins a la categorització semàntica.

Enumerarem els programes en forma de petites fitxes amb les característiques bàsiques, aturant-nos sobretot en la descripció i en les funcions més importants. Dels requisits de maquinari no en parlarem en detall: quasi tots els programes actuals necessiten màquines molt potents per a un bon rendiment. Diguem que han de ser ordinadors tipus Pentium 500 amb 64 RAM com a mínim.

<b>Nom:</b>	<b>ANNOTAPE</b>
<b>Distribució:</b>	Riverrun, Cork, Ireland.
<b>Sistemes operatius:</b>	Windows 98, WinNT 4.0., Mac OS x 10.2 o superior.
<b>Descripció:</b>	<p>Converteix l'ordinador en un magnetòfon virtual:</p> <p>Registra la parla, guarda els registres (fins a 100 hores), l'anota, l'etiqueta i transcriu els segments necessaris per a l'anàlisi.</p>
<b>Nom:</b>	<b>AQUAD Five 5.9.</b>
<b>Distribució:</b>	Verlag Inge Huber, Schwangau (Germany).
<b>Sistemes operatius:</b>	Win 9x, WinNT.
<b>Descripció:</b>	<p>Codificació del text en pantalla, en una sola fase; és a dir, no cal impressió i recodificació sobre la impressió. La codificació la fa totalment l'analista o es fa a partir de models de codificació del programa. També pot fer codificació en dues fases, a partir d'impressió. Mesures freqüencials. Diccionaris per a la categorització. Diferenciació de perfils d'autors o emissors.</p>
<b>Nom:</b>	<b>ATLAS. ti</b>
<b>Distribució:</b>	Sage Scolari.
<b>Sistemes operatius:</b>	Win 9x, WinNT.
<b>Descripció:</b>	<p>Processa textos i també dades gràfiques en àudio i vídeo. Codificació directa en pantalla. Definició de segments i de segments en context. Categorització semàntica. Compara segments textuais. Permet visualitzar mapes de relacions entre fragments o registres, i també entre categories.</p>
<b>Nom:</b>	<b>HYPERQUAL versió 11</b>
<b>Distribució:</b>	Hyperqual.
<b>Sistemes operatius:</b>	Mac OS 8.x o superior (Hypercard).
<b>Descripció:</b>	<p>És una eina específicament qualitativa. Treballa sobre transcripcions d'entrevistes i sobre documents. Ofereix un sistema classificador semblant a fitxes o targetes en què s'introdueixen les unitats registrades a les quals, en principi, no cal assignar codi. Després, a partir d'aquestes targetes, es pot fer la codificació i categorització.</p>
<b>Nom:</b>	<b>Hyperresearch 2.5.</b>
<b>Distribució:</b>	Sage, Inc.
<b>Sistemes operatius:</b>	Win 9x, WinNT, Mac OS 8.0, Power PPC.
<b>Descripció:</b>	<p>Permet treballar amb text, gràfics, àudio i vídeo. Té totes les funcions d'un cercador. Codificació en pantalla que, mitjançant un sistema expert, assigna automàticament un codi a cada ocurrència de mot o segment categorial. És capaç de fer codificacions múltiples. Gràcies a rutines d'intel·ligència artificial, es pot comprovar si les hipòtesis assumides se sostenen sobre la codificació final. També pot mostrar diagrames de relacions. Exporta les dades a paquets estadístics o fulls de càlcul.</p>
<b>Nom:</b>	<b>INTEXT 4.1. o TextQuest</b>
<b>Distribució:</b>	Harald Klein (autor).
<b>Sistemes operatius:</b>	DOS (INTEXT) o Windows (TextQuest).
<b>Descripció:</b>	<p>Fa possibles les categoritzacions automàtiques a partir de diccionaris del propi programa. Fa llistes de mots-en-context (KWIC) a partir de línies d'escriptura de longitud variable o bé de frases. Permet codificació interactiva. Hi ha una versió en castellà (TextQuest). Orientació bastant quantitativa.</p>

<b>Nom:</b>	<b>MaxQDA</b>
<b>Distribució:</b>	VERBI GmbH Anne Kuckartz.
<b>Sistemes operatius:</b>	Win 9x, WinNT (versions anglesa i alemanya).
<b>Descripció:</b>	Importa textos d'Internet, els organitza, selecciona segments de text com a unitats de registre i les codifica, crea sistemes de categories que poden ser jeràrquics, fins a deu nivells. Exporta les dades a diferents paquets estadístics o fulls de càlcul.

<b>Nom:</b>	<b>QSR NUD*IST (Non-numerical Unstructured Data Indexing Searching and Theorizing)</b>
<b>Distribució:</b>	Sage.
<b>Sistemes operatius:</b>	Win 9x, WinNT.
<b>Descripció:</b>	Útil sobretot per analitzar grans paquets de dades textuais o gravades, materials d'una certa complexitat. Completes funcions de cerca. Codificació en pantalla o automàtica (ràpida i rigorosa). Categorització, si cal, jeràrquica. Permet establir lligams entre dades quantitatives i qualitatives. Admet l'associació de contextos extratextuals a les categories. Treballa en connexió ( <i>on line</i> ) o fora d'aquesta ( <i>off line</i> ).

<b>Nom:</b>	<b>QSR NVIVO 2.0.</b>
<b>Distribució:</b>	Sage .
<b>Sistemes operatius:</b>	Win 9x, WinNT.
<b>Descripció:</b>	A més de tenir les funcions d'altres programes (cerca, computació), permet editar, codificar visualment i enllaçar documents mentre són creats, filtrats o treballats. Té un sistema de mòduls que brinda diferents vies de representació dels temes de recerca. Així, una sèrie de mots podria ser codificada en diversos mòduls que li fossin pertinents. Representa gràficament estructures relacionals. Ofereix sumaris estadístics.

<b>Nom:</b>	<b>Text Analyst 2.0.</b>
<b>Distribució:</b>	Megaputer.
<b>Sistemes operatius:</b>	Win 9x, WinNT.
<b>Descripció:</b>	Processador de textos intel·ligent capaç de realitzar anàlisis semàntiques, crear categories automàticament o fer resums i classificacions de textos o de parts de textos. Orientació bastant quantitativa.

<b>Nom:</b>	<b>Textbase alpha</b>
<b>Distribució:</b>	Bo Summerlund (autor).
<b>Sistemes operatius:</b>	DOS.
<b>Descripció:</b>	Especialitzat en anàlisi textual qualitativa en sentit estricte. Accepta el text directament tot i que està suportat per DOS. Si les dades ja estan estructurades <i>a priori</i> , com les d'una entrevista, aplica codificació automàtica; altrament, la codificació es fa en pantalla. Admet un gran nombre de superposicions i subordinacions o incrustacions. Permet exportar les dades a SPSS.

<b>Nom:</b>	<b>The Ethnograph 5. 04.</b>
<b>Distribució:</b>	Sage Scolari.
<b>Sistemes operatius:</b>	Win 3.1., Win 9x, WinNT.
<b>Descripció:</b>	Amb Nudist, un dels programes històricament més utilitzats, en ser dels pioners. De fet està orientat, en general, a dades recollides en qualsevol mena d'investigació qualitativa. Malgrat això és molt versàtil. Admet qualsevol processador de textos. Té les típiques funcions de cerca, codificació, comparació i establiment de relacions. Té gran capacitat i efectivitat (segments fins a 9.999 línies, dotze assignacions de codi per segment, jerarquies fins a set nivells, etc.).

## ■ Principals webs i informació a la Xarxa

- Podeu ampliar la llista anterior o alguns dels seus aspectes a partir dels webs següents:

<http://textanalysis.info/terms.htm>

Una pàgina que conté informació força completa sobre la majoria de programes actuals.

<http://caqdas.soc.surrey.ac.uk/>

Pàgina que proporciona informació, orientacions i ajuda pràctica (fins i tot entrenament) als usuaris de diferents programes d'anàlisi qualitativa. Hi ha demostracions gratuïtes.

<http://www.qual.auckland.ac.nz/>

Dóna igualment informació general sobre programes informàtics de suport a la recerca qualitativa.

- També podeu entrar directament en les pàgines d'aquests programes. Per exemple:

### **Annotape:**

<http://www.annotape.com>

### **Atlas ti:**

<http://www.atlasti.de/>

### **Nudist i NVIVO 2.0.:**

[http://www.qsrinternational.com/products/productoverview/product\\_overview.htm](http://www.qsrinternational.com/products/productoverview/product_overview.htm)

Conté demostracions gratuïtes.

### **HyperRESEARCH:**

<http://www.researchware.com>

No sols conté informació sobre el programa sinó també instruccions i consells útils.

### **The Ethnograph v 4.0.:**

<http://www.qualisresearch.com/>

Ofereix demostracions.

- Una altra via d'obtenció d'informació és entrar a la gran pàgina de la metodologia qualitativa, la de **The Qualitative Report:**

<http://www.nova.edu/ssss/QR/web.html>

Un cop a dins trobareu un llistat exhaustiu de pàgines relacionades amb la recerca qualitativa **a les quals podeu accedir simplement fent clic a l'entrada corresponent**. Totes les adreces que hem escrit aquí consten també en aquestes entrades:

### **Annotape**

<http://www.annotape.com>

### **The Atlas/ti Home Page**

<http://www.atlasti.de/>

### **Computer Assisted Qualitative Data Analysis Software (CAQDAS)**

<http://caqdas.soc.surrey.ac.uk/>

### **Harald Klein's Page of Textual Analysis Resources**

<http://textanalysis.info/terms.htm>

### **HyperRESEARCH**

<http://www.researchware.com>

### **Q.S.R. Qualitative Solutions of Research WWW Site**

[http://www.qsrinternational.com/products/productoverview/product\\_overview.htm](http://www.qsrinternational.com/products/productoverview/product_overview.htm)

### **/SWORLD**

<http://www.qual.auckland.ac.nz/>

### **The Ethnograph V4.0. Qualitative Research and Data Analysis Software**

<http://www.qualisresearch.com/>

A més a més, hi ha altres pàgines referents a programes que no hem inclòs en les fitxes anteriors com **AnSWR**

software o Qualrus.



■ Els principals programes d'anàlisi de contingut i qualitativa tenen pàgines a Internet força completes.

## Diagrama d'operacions bàsiques en una anàlisi de contingut

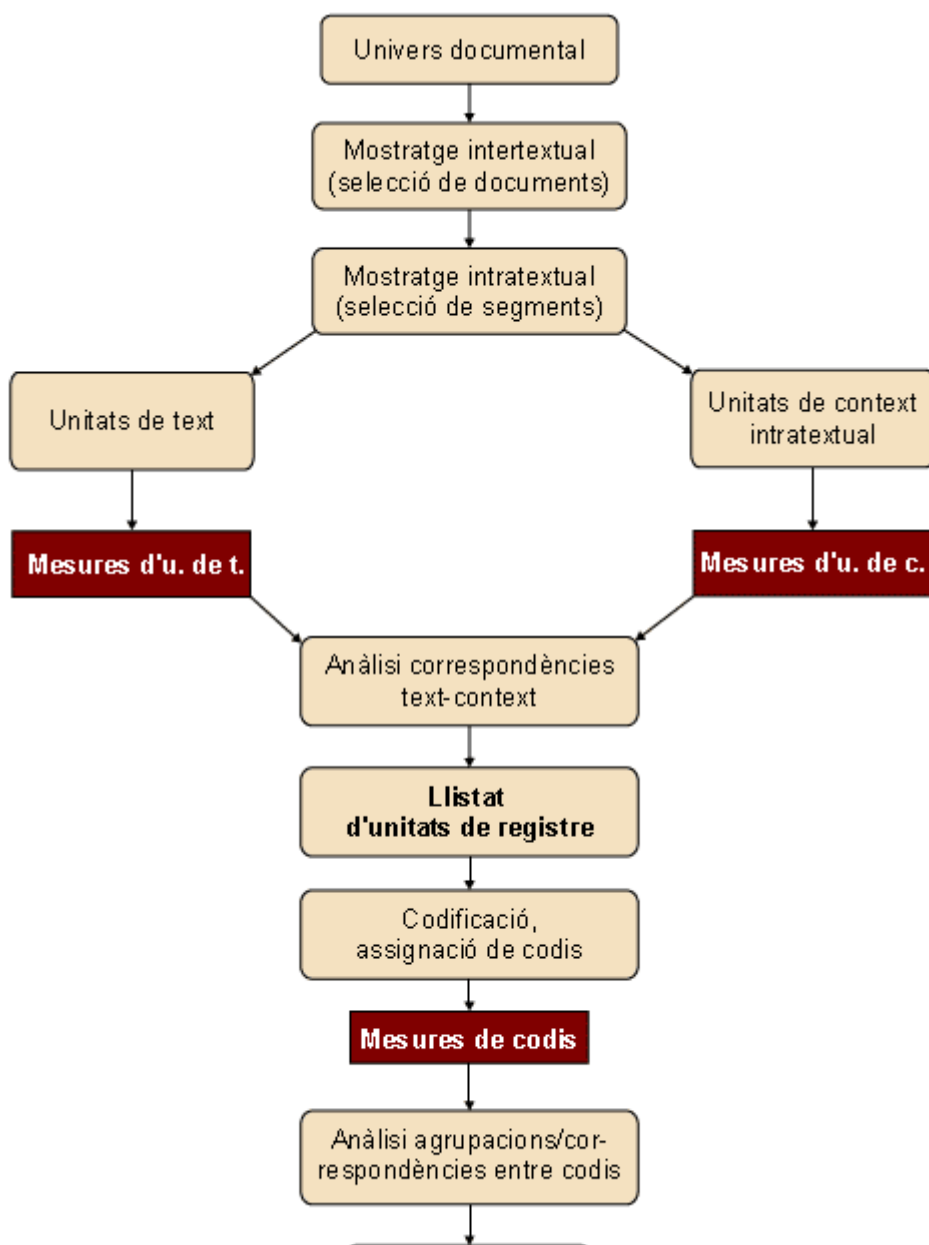


Una investigació d'anàlisi de contingut es pot representar mitjançant una seqüència d'operacions, i cada una de les quals pressuposa les anteriors.



En l'esquema adjunt us oferim la columna vertebral de la seqüència d'operacions i accions necessàries per dur a terme una anàlisi de contingut completa, des del seu punt de partida –l'establiment d'un univers documental– fins a les diferents direccions d'anàlisi que es poden seguir a propòsit de les dades obtingudes. Totes aquestes operacions han estat especificades en els apartats previs.

En el dibuix hem distingit amb dues tonalitats de color les operacions de mesura –per definició més "quantitatives"– de les operacions purament qualitatives (selecció, codificació, categorització, etc.).







**Les operacions necessàries per dur a terme una anàlisi de contingut mostren un paral·lelisme total amb les que exigeix qualsevol investigació observacional o qualitativa.**

### Exercicis d'autoavaluació

1. La peculiaritat dels sistemes de categories que classifiquen comportament comunicatiu o text resideix en el fet que...
  - no han de complir les condicions d'exclusivitat i exhaustivitat, mentre que els altres, sí.
  - compleixen la condició d'exclusivitat lògica entre categories de manera parcial mentre que els altres la compleixen de manera total.
  - poden complir la condició d'exclusivitat lògica entre categories gràcies a la combinació d'indicadors, mentre que els altres la compleixen sempre gràcies a l'exclusivitat entre indicadors.
  - compleixen la condició d'exhaustivitat lògica només pel que fa a categories, mentre que els altres la compleixen en tot el sistema.
  
2. Un corpus és el resultat de...
  - recollir i organitzar sistemàticament les dades produïdes per a una anàlisi de contingut.
  - fer la primera selecció dels materials que serviran per a fer un estudi, especialment de textos.
  - fer un mostratge de segments dins dels materials seleccionats en primera instància.
  - concentrar l'anàlisi sobre els segments més rellevants d'un text.
  
3. Una unitat de registre textual està composta per...
  - una unitat de text i una de context.
  - una unitat de text.
  - una unitat de context.
  - un mot i una unitat de context.
  
4. Quan interpretem una categoria respecte al seu context extratextual l'estem relacionant amb...
  - el context constituït per altres textos del corpus.
  - el context antecedent en el mateix text.
  - el context consegüent en el mateix text.
  - les condicions de producció del text (històriques, socials, biogràfiques).
  
5. Una anàlisi de contingut que examina i interpreta les seves categories en relació amb el destinatari del missatge o del text té una orientació semiòtica eminentment...
  - gramatical

- gramatical.
- semàntica.
- pragmàtica.
- sintàctica.

6. En donar la taxa o velocitat d'ocurrència d'un mot en un text, què seria millor, donar-la per pàgina o per paràgraf?

- per pàgina, ja que el nombre de mots per pàgina és fix.
- per paràgraf, ja que aquesta és una unitat sintàctica o estilística, no imposada per l'edició o la impressió.
- per paràgraf, ja que aquesta unitat té longitud variable.
- per pàgina, ja que aquesta és una unitat de grandària espacial fixa i relativament estable.

7. Una regla d'enumeració especifica sobretot...

- el marc de referència d'una mesura o l'àmbit espacial en què ha estat trobada.
- l'ordre en què s'han de considerar les categories a l'hora de ponderar.
- l'ordre seqüencial en què es presenten les categories en un text o secció de text.
- el context intratextual de la unitat de registre mesurada.

8. En l'anàlisi de contingut un codi és...

- la notació utilitzada per a una transcripció.
- la clau d'un document xifrat.
- els símbols utilitzats durant el registre per a recollir la informació d'un text.
- les etiquetes o noms posats a les diferents unitats de registre text-context.

9. En un text que comença a l'inici de la pàgina 14 i acaba a la meitat de la 27, la freqüència total d'una categoria X és  $f_x = 49$ . Quina és la taxa o velocitat d'ocurrència,  $v_x$ , de la categoria en qüestió?

- 3,5.
- 0,28/pàgina.
- 3,92/pàgina.
- 3,63/pàgina.

10. Al llarg d'un text hem registrat 12 vegades la categoria A ( $f_A = 12$ ), 29 vegades la B ( $f_B = 29$ ) i 113 vegades la C ( $f_C = 113$ ). Per tant, la freqüència relativa d'A,  $p_A$ , serà aproximadament...

- 0,08.
- 0,02.
- 12,83.
- 29.

11. Assenyaleu quin dels materials següents no es prestaria a una anàlisi de contingut.

- El missatge de final d'any del president, emès per la ràdio.
- Els diàlegs d'una pel·lícula.
- Les accions instrumentals d'un nen/a mentre dibuixa (agafar el llapis, aplicar-lo al paper, esborrar, etc.).
- L'etiqueta d'una ampolla de vi (reserva) explicant les excel·lències del producte.

12. Indiqueu, entre els recursos o tècniques següents, quina no es pot considerar dintre de l'anàlisi de contingut.

- Q-anàlisi.
- Taules de freqüència d'ús de mots segons longitud.
- Interpretació contextual seqüencial.
- Anàlisi conversacional.

13. En l'anàlisi de contingut el context intratextual s'entén habitualment...

- com l'oració en què està inserida una unitat de registre.
- com a mots, frases, oracions o paràgrafs a esquerra i dreta de la unitat considerada.
- com a mots, frases, oracions o paràgrafs al principi i final del text.
- com el paràgraf en què s'insereix una unitat de registre.

14. L'anàlisi de contingut va rebre un impuls decisiu gràcies...

- als estudis bíblics en el món protestant.
- a la crítica literària anglosaxona de la postguerra.
- a l'estudi de la publicitat visual en els seixanta.
- a l'estudi dels textos de propaganda nazi durant la Segona Guerra Mundial, per encàrrec del govern dels Estats Units.

15. Quan afirmem que un corpus és exhaustiu volem dir que...

- és complet en el nombre de pàgines necessàries, des d'un punt de vista estadístic, per a cobrir el material.
- comprèn tot el que ha escrit l'autor.
- recopila tot el que s'ha pogut trobar sobre el tema investigat fins al moment de l'estudi.
- cap dels elements previsibles sobre la base del criteri de selecció hi pot faltar.

16. Els materials estudiats en una anàlisi de contingut poden provenir de...

- la recopilació de documents ja existents i producció expressa de documents nous (però no de transcripcions).
- la recopilació de transcripcions ja existents i producció expressa de transcripcions noves (però no de documents).
- la recopilació de documents i transcripcions ja existents i producció expressa de documents i transcripcions noves.
- la recopilació de documents i transcripcions ja existents, però no producció de noves.

17. Què distingeix, en la pràctica, l'enfocament qualitatiu del quantitatiu en l'anàlisi de contingut?

- Senzillament la no-utilització de mesures i anàlisis estadístiques.
- L'ús d'estadística no paramètrica en comptes de paramètrica.
- El fet d'utilitzar mesures només de tipus categorial o ordinal.
- El fet de derivar les inferències de la presència/absència d'un tret més que de la seva freqüència.

18. L'expressió  $\hat{\theta} = 1 - \frac{N_1}{T}$ , on  $N_1$  és el nombre de categories detectades només un cop i  $T$  el nombre total d'unitats de registre recollides, serveix per a...

- aturar el mostratge de segments en un text quan l'expressió s'acosta a 1.
- aturar el mostratge de segments en un text quan l'expressió s'acosta a 0.
- fixar el nombre de mots que ha de tenir el text quan l'expressió és igual a 0.
- fixar la grandària del sistema de categories utilitzat, és a dir, el nombre de categories que té el sistema.

19. La unitat textual de delimitació més difícil és...

- el paràgraf.
- la frase feta.
- el tema.
- l'oració.

20. Com que hi ha correspondència entre la grandària de les unitats textuales i la de les contextuales el context intratextual d'una oració serà...

- més gran que el d'un paràgraf.
- més petit que el d'un mot.
- més gran que el d'un tema.
- més gran que el d'un mot.

21. Un indicador o una unitat de registre estan relacionats amb la categoria o categories que els pertocuen mitjançant...

- una variable intermediària.
- una mesura (per exemple, "llavis caiguts" mesura tristesa).
- una atribució de significat (per exemple, "llavis caiguts" significa tristesa).
- una atribució de trets físics o externs ("llavis caiguts" és descrit per mitjà dels trets facials, de la pell, de la musculatura, etc.).

22. Un sistema de categories disponible, estandarditzat i implementat en un programa informàtic d'anàlisi de contingut és denominat...

- enciclopèdia.
- banc de dades textuales.
- diccionari o índex.
- corpus digital.

23. Supposeu que l'anàlisi de la publicitat en un diari es fa sobre la base d'un sistema de categories compost per tres grans categories: "Espectacles i esports", "Béns de consum" i "Habitatge". Quin tipus de criteri de classificació ha estat emprat des del punt de vista semiòtic?

- Sintàctica.
- Semàntica.
- Pragmàtica.
- Lingüística.

24. Les mesures de direcció se solen resoldre mitjançant...

- correlacions.
- anàlisi seqüencial.
- qüestionaris amb escales d'estimació i els índexs resultants.
- escales físiques i topogràfiques.
25. Per tal de treballar amb una certa profunditat teòrica i de continguts un programa informàtic d'anàlisi textual ha de ser capaç...
- de recuperar text (*text retrievers*), almenys amb la funció de cerca.
- de marcar els segments registrables mostrejats.
- d'assignar un contingut abstracte o semàntic als segments o a les unitats marcats.
- de detectar coocurrències dintre de contextos intratextuals.
26. En un sistema de categories textual les condicions d'exhaustivitat i exclusivitat temporal que ha de satisfer el sistema no poden ser mai preses en consideració.
- Cert
- Fals
27. Una anàlisi de contingut amb orientació semàntica se centra en l'ordre i organització amb què es presenten les categories al llarg del text.
- Cert
- Fals
28. En l'anàlisi de les respostes d'una entrevista estructurada, amb qüestionari, la categorització ja pot venir donada pel sistema que formen les mateixes respostes previstes, tancades o amb alternativa múltiple.
- Cert
- Fals
29. Una metodologia autènticament qualitativa hauria de rebutjar l'ús de programes d'ordinador per a l'anàlisi textual.
- Cert
- Fals
30. Una opció vàlida de mostratge intratextual és no fer-lo, és a dir, estudiar la totalitat del corpus textual reunit, sempre que aquest no sigui prohibitivament extens.
- Cert
- Fals

## Enllaços d'interès

A més de les pàgines web ja proporcionades en la part del text corresponent, podeu consultar:

<http://cbae.nmsu.edu/~dboje/qm>

Pàgina dedicada a l'anàlisi de material de comunicació en general, amb un punt de mira més ampli que el de l'anàlisi de contingut.

<http://www.asanet.org/sectionemca/LINKS.html>

Ofereix informació i idees sobre anàlisi de textos des del punt de vista etnometodològic i conversacional, amb adreces d'altres pàgines que tenen idèntica orientació.

<http://www.conversation-analysis.net/contents.htm>,

També centrada en anàlisi des del punt de vista conversacional.

## **Bibliografia**

Bardin, L. (1986). *Análisis de contenido*. Madrid: Akal.

González Rey, F. (1997). *Epistemología cualitativa y subjetividad*. La Habana (Cuba): Pueblo y Educación. [El text recomanat és l'últim capítol del llibre, "El análisis de contenido en la utilización de instrumentos cualitativos", pp. 193-280. És un text difícil però dóna oportunitat a l'estudiant de tenir contacte amb anàlisis totalment qualitatives i centrades en qüestions psicològiques.]

Krippendorff, K. (1980). *Content Analysis: An Introduction to its Methodology*. Beverly Hills (Cal.): Sage. [És un text molt citat però té un tarannà força quantitatiu.]

Weitzman, E. A. i Miles, M. B. (1995). *Computer Programs for Qualitative Data Analysis*. Thousand Oaks (Cal.): Sage.