

Katia Flucha Martinez

LA CONCORDANÇA DE GÈNERE SEMÀNTIC EN CATALÀ

TREBALL DE FI DE GRAU

dirigit pel Sr. Josep Demestre Viladevall

Grau de Psicologia



UNIVERSITAT ROVIRA I VIRGILI

Tarragona

2024

ÍNDEX:

1. Introducció.....	4
2. Objectius.....	20
3. Metodologia.....	21
4. Resultats.....	24
5. Discussió i conclusions.....	26
6. Bibliografia.....	28
7. Apèndix.....	30

Resum

Vam dur a terme un experiment d'eye-tracking sobre el processament de frases en català amb l'objectiu d'investigar l'efecte que té la no concordança de gènere i possibles diferències quan aquesta violació es dona amb un nom propi o un nom comú. Quan el parser es troba una no concordança de gènere, inicia un procés de *repair*, tenint dos possibles vies per a la condició de nom comú (reparar l'article o el nom), mentre que només en tindria una per la condició de nom propi (reparar l'article). Per tant, el que seria d'esperar és que la condició de nom comú i no-concordança tinguin temps de resposta més alts. En els nostres resultats vam poder confirmar aquests efectes, tenint temps de lectura més alts en la condició de nom comú sobre la de nom propi, i la condició de no-concordança sobre la de concordança. A més a més, vam trobar en la regió post-crítica una diferència significativa entre la condició de no-concordança sobre la condició de concordança únicament per als noms comuns, mentre que aquesta diferència no era significativa per als noms propis.

Abstract

We conducted a sentence processing experiment in Catalan using eye-tracking, to investigate the effect of gender disagreement and possible differences when this violation occurs with a proper noun or a common noun. When the parser finds a gender disagreement, it initiates a repair process, having two possible ways for the common noun condition (repairing the article or the noun), while it would only have one for the proper noun condition (repairing the article). Therefore, what would be expected is for the common noun disagreement condition and the disagreement condition to have higher response times. In our results, we were able to confirm these effects, finding higher reading times for the common noun condition over the proper noun condition, and for the disagreement condition over the agreement condition. Moreover, we found in the post-critical region a significant difference between the disagreement condition over the agreement condition only for common nouns, while this difference was not significant for proper nouns.

1. Introducció

Des del naixement, el llenguatge forma part del nostre dia a dia i, tot i que no sol requerir esforç i ser aparentment automàtic, és un procés molt complex que alhora és la base d'activitats individuals, socials, culturals i intel·lectuals i sobre el que encara tenim moltes incògnites. Com a humans, tenim la capacitat de transmetre qualsevol missatge que puguem imaginar, creant i entenent frases noves a diari seguint únicament unes normes gramaticals apreses i el vocabulari del que disposem. Tant per transmetre una idea com per rebre-la de forma que concordi amb el missatge inicial, molts processos hi intervenen, des de crear representacions de significat i de la frase amb el lexicon a crear representacions fonològiques per la persona que vol transmetre la informació (Fernández & Smith Cairns, 2010). La persona receptora del missatge haurà d'utilitzar les regles de la sintaxi i el seu vocabulari per avaluar el significat de cada paraula i la frase en general. Al sistema humà que fa aquesta segona part del procés se l'anomena *parser* i és el que s'encarrega del processament sintàctic (o *parsing*). Durant aquest processament es construeix una representació mental tenint en compte tant l'estructura de l'oració i les relacions (o dependències) que hi ha entre els diferents elements (paraules, sintagmes) que formen part d'aquesta, com les regles gramaticals. El que es busca entendre en els estudis de processament de frases és com el parser converteix en temps real una cadena de paraules en una representació estructural mental de la frase, quan el parlant i l'oient construeixen estructures sintàctiques utilitzant fonts d'informació diferents.

El processament de frases implica que els oients i lectors computen l'estructura sintàctica basant-se en el seu coneixement del llenguatge, ja que la senyal en si no conté informació sintàctica i per fer-ho utilitza clàusules. Una clàusula consisteix en un verb i els seus arguments, representant una unitat integrada de significat i estructura, el que les fa adequades per al processament. Les clàusules són manejables per al emmagatzematge en la memòria de treball durant el processament. Les frases amb clàusules amb límits ben marcats tenen un menor cost de processament que aquelles amb límits més difusos perquè les primeres porten a una menor probabilitat que la frase sigui ambigua. El parser, a més a més de descompondre la senyal en clàusules i computar la seva estructura interna, ha de fer front a ambigüitats estructurals. Un fet interessant

vist en estudis amb frases ambigües és que, les persones normalment no perceben l'ambigüitat global i tenen preferència per una interpretació concreta. Pel que fa a la reconstrucció de l'estructura d'una frase, el parser ha de computar ràpidament les relacions jeràrquiques entre paraules que li arriben en una seqüència lineal, enfrontant-se a múltiples opcions estructurals. Utilitza la gramàtica (sintaxi i morfologia) per identificar noms i verbs, la informació lèxica per obtenir els possibles marcs estructurals, i segueix rutines per minimitzar els recursos computacionals i optant per estructures simples, com proposa el model de processament de frases *garden path*. Aquest tipus de frases et porten a fer una primera interpretació incorrecta, per després fer que ens adonem que no li estem trobant el sentit, i finalment a re-analitzar quina ruta podria portar o no a la interpretació correcta. Hi ha tres possibles estratègies que utilitza el parser per resoldre les ambigüitats: *minimal attachment*, *late closure* i *active gap filling*.

L'estratègia de *minimal attachment* incorpora el material lèxic per construir l'estructura més simple possible. Si aquesta construcció és incorrecta, el parser ha de tornar a fer l'anàlisi i, per tant, tindrà un cost de processament més alt. D'altra banda, el cost de processament és baix si l'estructura mínima és correcta o si hi ha indicis per evitar errors. Això implica que en frases amb ambigüitat, la interpretació que es preferirà serà la que inclogui l'estructura mínima. Si el parser es troba, en intentar resoldre una ambigüitat local, que no hi ha una alternativa que la tingui (i son igualment complexes), aplicarà l'estratègia de *late closure*. En aquesta, s'adjunta el material nou al constituent més recent, com es veu en l'exemple "Algú va disparar a la criada de l'actriu que estava al balcó", on la clàusula "que estava al balcó" es ambigua (podent referir-se tant a l'actriu com a la criada). El *late closure* prediu que es preferirà modificar el nom més recent, en aquest cas, l'actriu tot i que alguns estudis han trobat que aquesta estratègia no és universal, havent-hi situacions en les que no es segueix aquesta per certes llengües.

En sintaxi, a més a més de moure elements dins d'una frase, es crucial identificar els forats (o *gaps*) deixats per aquests moviments. L'estratègia *active gap filling* es veu en frases com "Which car did Mike force Mary to buy?", on el processament és més costós que en altres frases ja que el parser ha de buscar un lloc on posar el *filler*, essent obstaculitzat per la presència d'altres elements en la posició potencial del *gap* (en aquest cas, Mary ocupa el lloc d'un primer

possible gap i el parser ha de seguir buscant fins al final de la frase). Relacionat amb aquest últim efecte, el parser també ha d'encarregar-se d'identificar les frases nominals a les que es refereixen els pronoms en una frase. Un exemple d'això es veu fàcilment en anglès on els pronoms reflexius, com "himself", han de referir-se a un nom en la mateixa clàusula, mentre que els pronoms personals com "him", han de referir-se a un nom fora de la seva clàusula:

- En "The boxer told the skier that the doctor for the team would blame himself for the recent injury", "himself" ha de referir-se a "doctor" perquè està en la mateixa clàusula.
- En "The boxer told the skier that the doctor for the team would blame him for the recent injury", "him" ha de referir-se a "boxer" o "skier" perquè "doctor" està en la mateixa clàusula.

També s'han fet estudis per veure si el coneixement del món real redueix l'efecte de *garden path* en frases com "The performer sent the flowers was very pleased" i "The florist sent the flowers was very pleased". Tot i que s'esperava que el coneixement de que els artistes reben flors reduiria l'efecte de la primera frase, ambdós casos van mostrar forts efectes de *garden path* en el seguiment ocular. Això ens indica que el parser no va tenir en compte el coneixement del món real en el seu anàlisi inicial. Aquest coneixement, però, va ajudar als participants a parafrasejar les frases, essent més fàcil amb les oracions sobre artistes. Això suggereix que el coneixement del món real és útil, però difícil d'aplicar en frases aïllades sense context (Fernández & Smith Cairns, 2010). Durant aquest procés, el parser també haurà de descodificar i integrar senyals de gènere i concordança, buscant també la forma més econòmica de reparar les frases que ho necessitin.

En els últims anys hi ha hagut força estudis sobre aquests processos en diferents llengües i utilitzant diferents mètodes experimentals, però encara destaca la manca d'estudis centrats en el català.

Gender cues i concordança de gènere

La concordança de gènere és un procés en el que es formen relacions sintàctiques entre elements, després de valorar el gènere del nom (Popov & Bastiaanse, 2018). Hi ha dos formes per a fer aquesta valoració: o bé deduint-ho a partir d'un element dependent o receptor de concordança (al que anomenarem *probe*) o a través directament del nom o pronom (al qual ens

referirem com a *goal*). La diferència entre aquests dos és que únicament el *goal* està marcat lèxicament amb el gènere (és la paraula de la qual depenen les altres) mentre que aquest valor es copia en el *probe* en el procés de concordança. La característica lèxica del gènere del *goal* es pot representar en diferents nivells de representació com: informació sintàctica (lema), informació morfològica (morfema de gènere) o informació semàntica (Vigliocco & Franck, 1999).

El gènere sintàctic o gramatical no aporta informació semàntica o pragmàtica addicional al context i per tant, no es pot deduir si és masculí o femení basant-se en la semàntica. En aquest cas podríem fer una deducció mirant la flexió de gènere del *probe* que és la única marca de gènere que tindríem disponible (Popov et al., 2020).

A nivell morfològic, diem que les característiques de concordança de gènere s'organitzen seguint una jerarquia de marcatge, que consisteix en que existeix una asimetria en la oposició masculí-femení, que fa que un dels dos sigui el gènere per defecte (i per tant sigui l'estatus predeterminat de la paraula) mentre que l'altre si que la té. En l'Espanyol, els noms poden ser o bé masculins o femenins i veiem que quan són masculins no solen estar marcats amb elements de gènere i poden modificar elements sense gènere o femenins, mentre que el valor femení es restringeix als noms femenins i si que està marcat. Un exemple d'això es pot observar en la frase "*demasiados paras en ese párrafo*", on un element que pot tenir tant una forma masculina com femenina (en aquest cas "*demasiados*"), només pot agafar la forma masculina en interaccionar amb una paraula sense gènere inherent com és la preposició "*para*". També quan s'uneixen noms de diferents gèneres, tots els elements que es poden modificar per crear la concordança de gènere han de prendre la forma masculina (Aleman Banon & Rothman, 2016).

També hem de tenir en compte que hi ha noms que diem que son transparents, en el sentit que el seu final representa de forma vàlida una marca de gènere, i noms opacs que no ens donen informació del gènere gramatical per la seva terminació. En Espanyol (i en llengües romàniques en general) hi ha generalment una forta correlació entre la terminació d'un nom i el gènere al que pertany, essent femenina la majoria dels que terminen en "-a" i masculins els que ho fan en "-o", i el que fan que siguin noms transparents (Caffarra et al., 2015). Això també ho podem veure en el Italià, tenint la mateixa terminació els

noms de "-a" i "-o" i en el cas dels opacs, en aquest cas terminen habitualment en "-e". A més a més, quan el gènere ve donat biològicament, l'arrel de la paraula sol ser diferent (com per exemple en madr-e -padr-e) mentre que en els transparents amb gènere semàntic normalment comparteixen arrel (com en bambin-o/a) (Popov et al., 2020).

En el cas del català, la forma més predominant en els noms és que els masculins no tinguin marcatge de gènere mentre que els femenins sí. Això és, un nom masculí com podria ser "amic" no presenta transparència (seria opac), a diferència del l'espanyol o l'italià, on habitualment veuríem una terminació en "o", i només la seva oposició en femení presentaria la marca de gènere "a" (com en "amiga").

D'altra banda, tenim el gènere semàntic (o biològic) que fa referència al grup de noms (usualment referits a persones, professions i alguns animals) que tenen un referent al món amb sexe biològic, el qual es sol preservar en el llenguatge, fent que si un referent es masculí en el món, també serà masculí el seu nom. És el més comú que es mantingui el mateix gènere en el lèxic que en el biològic, el que fa que aquest tipus de marca de gènere sigui molt fiable (Popov et al., 2020). En aquest estudi ens centrem en el gènere semàntic, més concretament en els noms comuns i els noms propis. En el cas del català, veiem que en aquests dos tipus de noms tornen a seguir la forma explicada anteriorment, on només els noms femenins acostumen a tenir una marca de gènere ("-a") mentre que els masculins no la tenen.

El sistema de gènere en català

En el cas del català, els noms no tenen marques de gènere uniformes associades a cada gènere, sinó que varien segons el nom. Hi ha diferents classes formals, i en el cas dels noms masculins són les següents:

La classe que trobem en la majoria dels casos és pels noms que no tenen cap marca flexiva i, per tant, es redueixen al radical (com per exemple amor, camió o cap). Normalment aquests acaben en consonant, sent excepcions els que acaben en vocal (com seria geni, odi o sofà).

En segon lloc, trobem noms que acaben en -e. Aquesta vocal, en la majoria dels casos, fa la funció de suport, per a que aquestes paraules no acabin en una seqüència de consonants (arbre, batlle, deixeble), però també hi ha contextos en que no està condicionada fonològicament i fa una funció de marca flexiva (pare, període, hoste).

Menys freqüentment trobem els noms que afegeixen la -o al radical (amo, carro, musclo) o la -a, característica dels noms femenins (dia, mapa, papa).

Un grup petit, normalment referit a la procedència ètnica o geogràfica, són els que afegeixen -u en paraules que acaben amb el diftong -eu (europeu, hebreu).

Molt esporàdicament (normalment per paraules que el català ha agafat d'una altra llengua, és a dir, manlleus), els noms masculins poden tenir altres marques més idiosincràtiques i difícilment sistematitzables (anus, virus, tipus).

Pel que fa als noms femenins, trobem també diferents classes:

El més habitual i regular presenta la marca -a (casa, dona, porta).

De forma molt menys nombrosa, estan els noms femenins sense cap marca flexiva (com col, dent o sal), els derivats amb sufixos -ció (organització), -ió (divisió), -itat (cordialitat), qualsevol variant de -dat, -edat, -tat i -etat (bondat, claredat, lleialtat, seguretat), -etud o -itud (inquietud, solitud), i la majoria dels que tenen el sufix -or (calor).

Molt esporàdicament, en paraules que provenen de manlleus, podem trobar terminacions en -o (libido), -e (sèrie), i molt puntualment, en -us.

Quan els noms no tenen marcatge de gènere, l'oposició es pot aconseguir per procediments de concordança, per mots de radical o per derivatius:

Quan es fa a través de concordança, és habitual que sigui amb l'article (el cantant/la cantant).

Quan es tracta de radicals diferents, utilitzen paraules diferents amb sufixos característics de cada gènere (home/dona, hereu/pubilla).

Per últim, en alguns noms s'estableix l'oposició de gènere mitjançant sufixos derivatius o terminacions diferents. Es pot formar el nom femení afegint un sufix al radical masculí, com seria -ina o -essa (rei/reina, alcalde/alcaldeessa), essent d'aquests dos el més prevalent -essa, tot i que avui en dia s'opta pel sufix -a (president/presidenta). Es pot formar el nom masculí afegint el sufix -ot al radical femení (nina/ninot), tot i que aquesta es dona en un grup petit de noms. També hi ha noms en que el masculí i el femení tenen terminacions diferents, amb els sufixos -dor/-driu o -leg/-loga (actor/actriu, psicòleg/psicòloga).

Pel que fa al marcatge de gènere en noms que designen entitats inanimades, en la majoria dels casos no hi ha cap raó aparent que justifiqui que un nom tingui un gènere o l'altre i aquesta distinció no és sistemàtica quan el radical és el mateix però amb marques de gènere diferents.

Veiem també la manca de raó aparent en la distinció de gènere per entitats inanimades amb els noms que tenen doble gènere, és a dir, la paraula és la mateixa però amb significat diferent en funció del gènere (com es exemplificat en la Taula 1). En alguns casos, el significat també pot ser el mateix, però amb una associació a contextos diferents en funció de l'oposició.

Masculí	Femení
El clau (que es clava)	La clau (que obre i tanca un pany)
El còlera (malaltia)	La còlera (indignació)
El fi (objectiu, finalitat)	La fi (acabament).
El mar.	La mar.

Taula 1. Exemples de noms amb doble gènere

Hi ha un grup de noms que són invariables (és a dir, que no tenen l'alternança -e/-a com és el cas de perfecte/perfecta o culte/culta) acabats en -e, que, com hem dit abans, té la funció de vocal de suport en la majoria del casos. En la Taula 2, veiem alguns exemples (també essent invariables altres noms com enorme, immune, lliure, etc.).

<i>Terminació</i>	<i>Exemples</i>
-ble	amable, criticable, feble, increïble, notable, possible
-aire	preguntaire, rondinaire, somicaire, trinxaire, xerraire
-ense	castrense, circense, forense, hortense, tarraconense
-oide	androide, antroipoide, cel·luloide
-ànim	equànim, exànim, pusil·lànim, unànim
-forme	conforme, deforme, oviforme, uniforme
-angle	acutangle, rectangle, triangle
-corde	concorde, discorde
-corne	bicorne, tricorne, unicorne
-lingüe	bilingüe, monolingüe, plurilingüe
consonant + -re	alegre, bímembre, cafre, cèlebre, fúnebre, grandirostre, il·lustre, lacustre, lúgubre, mediocre, terrestre (però aspre/aspra, destre/destra, íntegre/íntegra, negre/negra, pobre/pobra, pulcre/pulcra, tendre/tendra)

Taula 2. Principals terminacions en -e dels adjectius invariables

Un petit grup de noms prenen el sufix -a per a ambdós gènere, normalment en els adjectius que poden funcionar també com a noms (homicida, pirata, indígena).

Per últim, quan l'oposició de gènere es dona per l'alternança -ø/-a, veiem canvis també en el radical al que afecten. Aquests canvis normalment afecten a la consonant final del radical de la forma masculina, com afegir la consonant n

(cosí/cosina), fer la r final sonora (dur/dura), ensordir els obstruents (amic/amiga), canviar al u per v (blau/blava) i la l per l·l (tranquil/tranquil·la.

Descodificació de gènere i mecanismes de repair (reparació)

Popov et al. (2020) diuen que en el procés de descodificació de gènere en lectura, quan el nom és transparent, com és l'exemple en italià de *treno* (tren) que va precedit de *il* (el) (veure la figura 1), tenen lloc dos possibles processos simultanis: el que el parser accedeixi al lema del nom a través de l'arrel únicament, ja que aquest li dona la informació suficient; i es processa alhora com a masculí el marcador de gènere "-o", reafirmant el valor de la característica de gènere del nom. En aquest cas es faria amb èxit la recuperació de gènere ja que hi ha concordança i es podria seguir amb el processament sense obstacles.

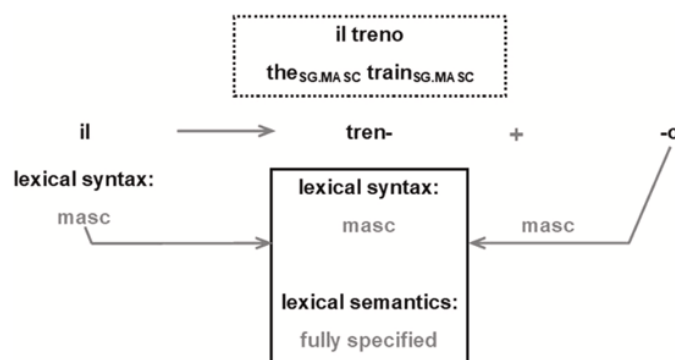


Figura 1. Esquema del mecanisme de processament visual sintàctic de gènere assumit

Quan l'arrel del nom és opaca (com es veu en la figura 2), el parser accedeix a l'article masculí *il* primer, conservant-lo per comparar-lo amb el *goal*. El parser ha de confiar completament en el morfema de gènere "-o" ja que en aquest cas no pot utilitzar l'arrel "babin-" perquè no li dona cap informació ni de sintaxi lèxica ni de semàntica, però si el nom va precedit d'un *probe*, assumirà que el gènere del *goal* coincidirà amb el del *probe*. Per tant, en aquest cas, la especificació de gènere prové de l'element receptor i de la informació de gènere del morfema del nom. En cas que la informació de gènere del *probe* i la del morfema de gènere no coincideixin hi ha una competència per especificar el valor de gènere del lema, tant en termes de sintaxi lèxica com de semàntica lèxica (Popov et al., 2020).

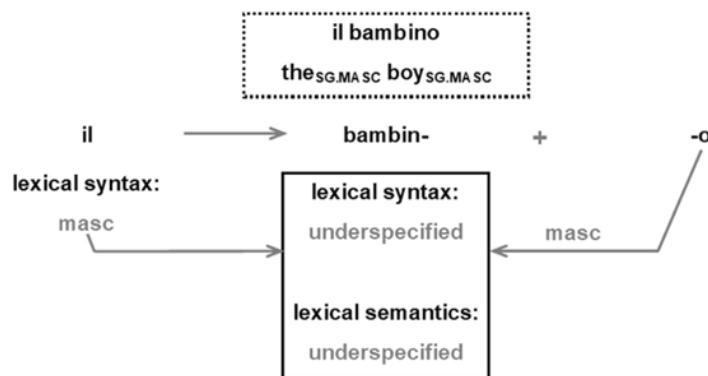


Figura 2. Esquema del mecanisme de processament visual sintàctic de gènere assumit

En cas que no hi hagi concordança de gènere entre l'article i el nom, es donarien processos de reparació. Popov, et al. (2020), en l'estudi sobre l'italià, proposen dues formes diferents de reparació (o *repair*) en funció de si el nom té gènere sintàctic o gènere semàntic.

Si el substantiu té gènere sintàctic el parser té una única via directa per reparar la no concordança (exemplificat en la figura 3). En el cas de "la treno", tant l'arrel (tren-) com el sufix (-o) estableixen que el nom és masculí, per tant, l'única via de reparació seria canviar l'article femení *la* per un masculí (*il*) que seria el correcte en aquest context. Els noms amb gènere sintàctic es consideren més fàcils de reparar, no només pel fet que només existeix una via de reparació possible, sinó perquè es reparen a nivell morfosintàctic, ja que només implica la recuperació de gènere a partir del lema (arrel) i la morfologia (terminació).

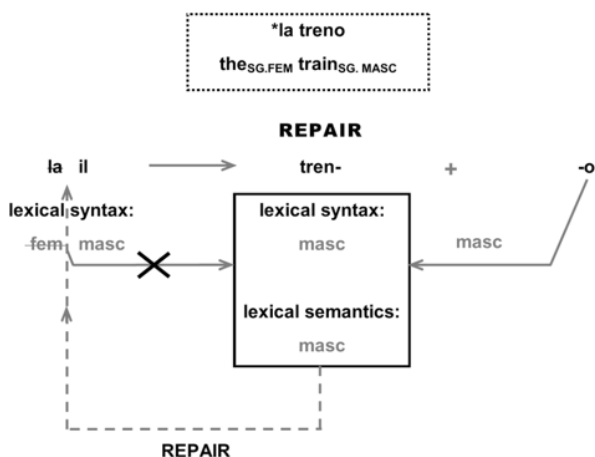


Figura 3. Mecanismes de repair per noms amb gènere sintàctic

Aquest mateix cas en el que hi ha una única via de reparació el veuríem en el cas de les no concordançes de gènere amb noms propis en català que nomen tenen o bé la forma femenina o la masculina (com els que utilitzem en el nostre estudi). Per tant, aquests noms no permeten que es pugui canviar el seu gènere

afegint un sufix. Per exemple, *Antoni* si que podria convertir-se en un nom femení, ja que *Antonia* és un nom que també s'utilitza i, en canvi, *Marc* no es podria canviar (ja que no existeix el nom *Marca*). L'única via per crear concordança de gènere amb el cas de la figura 4 "La Marc" seria canviar el gènere de l'article de femení a masculí, ja que el nom propi *Marc* no té un equivalent en femení canviant el sufix d'aquest.

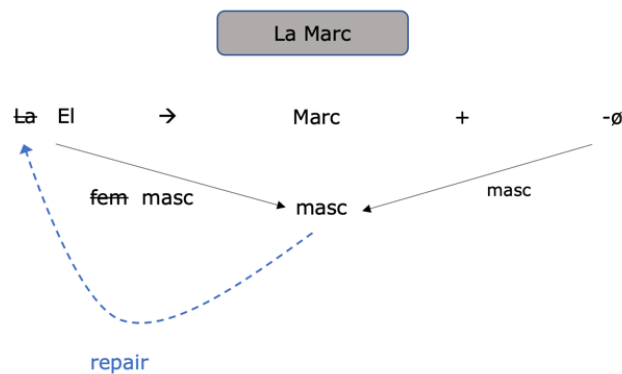


Figura 4. Mecanismes de repair per noms propis en català

D'altra banda, els processos de reparació amb noms amb gènere semàntic són més complexos ja que hi ha dues possibles vies de reparació. La complexitat ve també donada per la subespecificació tant sintàctica com semàntica de l'arrel, ja que aquesta necessita el valor de gènere per especificar el sexe biològic del referent (és a dir, la semàntica de l'arrel és neutra pel que fa al gènere i és el morfema de gènere el que ens indica si és mascle o femella). Per tant, el tret de gènere s'ha de recuperar a través del sufix que, no només avaluarà el gènere del nom, sinó que també especificarà si el referent animat és mascle o femella. En aquesta situació doncs, com queda exemplificat en la figura 5, hi haurà dues vies de reparació igual de plausibles. En la primera el parser accepta com a correcte el valor de gènere del sufix del nom (-o masculí) i repara l'article convertint-lo del femení *la* a masculí *il*. En la segona, s'assumeix que el correcte és el valor de gènere de l'article, canviant el sufix del substantiu de *bambin-o* (nen) a *bambin-a* (nena). Així que s'assumeix que la reparació de gènere en noms amb gènere biològic es més complexa no només perquè el número de vies de reparació és més alt, sinó també perquè opera tant a nivell morfosintàctic com a nivell semàntic. Per tant, qualsevol no concordança de gènere amb noms semàntics crearia una no concordança, perquè la especificació del sexe del referent (semàntica lèxica) depèn del valor del gènere (sintaxi lèxica). Com a conseqüència, el procés de reparació no només afectaria el tret de gènere

(sintaxi lèxica) sinó que també hi ha hauria una reparació o reanàlisi de la semàntica lèxica (Popov et al., 2020).

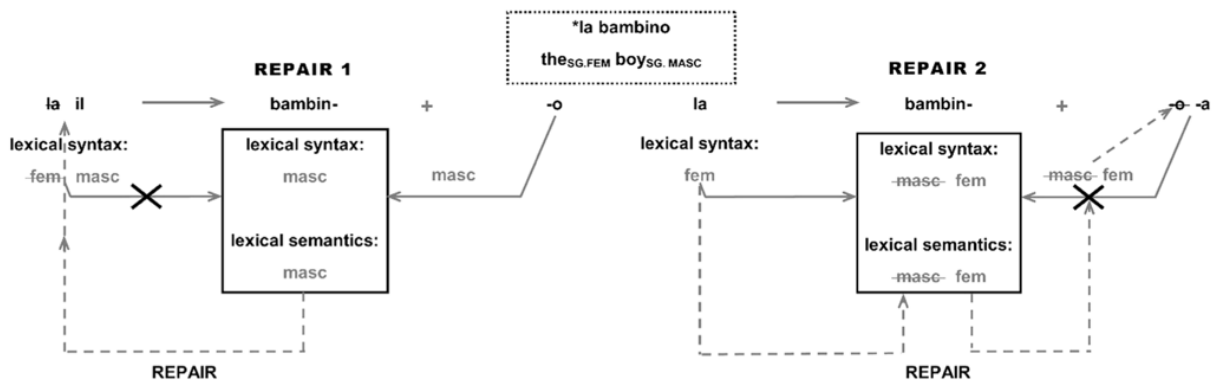


Figura 5. Mecanismes de repair per noms amb gènere semàntic

Quan el parser es troba amb una no concordança de gènere en català amb un nom comú també té dues possibles vies de reparació. En el primer (representat en la figura 6) veiem que accepta com a correcte l'absència de marca de gènere en el nom, que ens indicaria que està en masculí, i per tant repara l'article. En la segona en canvi, assumeix que el que està correcte és el gènere que veu en l'article, reparant doncs l'arrel de la paraula (la sintaxi i semàntica lèxica) i afegint el sufix "-a" per fer-lo també femení.

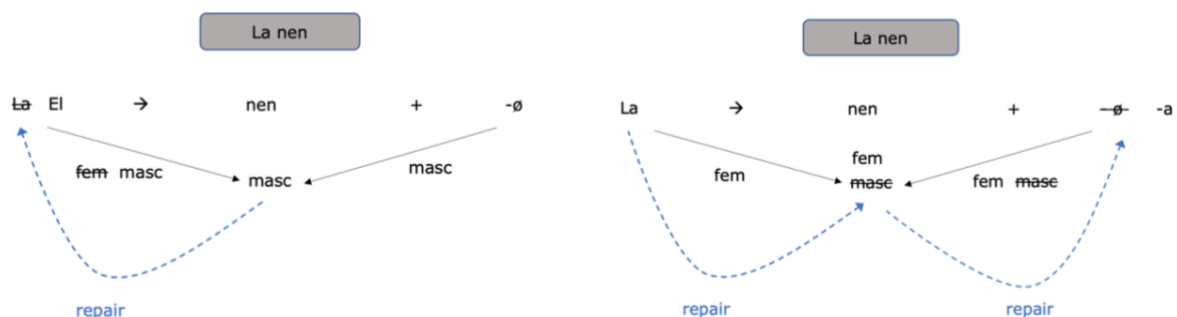


Figura 6. Mecanismes de repair per noms comuns en català

Abans d'entrar en els detalls de l'estudi, parlarem de diferents tècniques on-line que s'utilitzen en els estudis sobre el processament d'oracions.

Mètodes experimentals en processament de frases

Fa relativament poc, els experiments en psicolingüística van passar de fer-se responnent qüestionaris amb paper i llapis (el que es diu mètodes *offline*) a mètodes *online* que ens permeten obtenir informació sobre la interpretació de frases mentre està essent llegida o escoltada cada paraula. Aquesta nova forma de dur a terme els experiments també ens aporta possiblement puntuacions més

congruents amb el coneixement implícit del participant, minimitzant per tant l'ús de coneixements apresos explícitament, ja que mesurem interpretacions en temps real. Son tres els mètodes més freqüentment utilitzats:

En el *self-paced reading* es presenten paraules o segments de frases cada cop que es clica una tecla designada. S'analitza el temps que triga cada participant en llegir cada segment presentat, eliminant les puntuacions de menys de entre 100-200 ms, ja que es considera que, en tan poc temps, no ens dona temps de computar la informació el suficient com per a que aquesta puntuació sigui informativa pel nostre experiment. També s'eliminen temps massa alts (que es pot determinar perquè supera els 2000-6000 ms o per altres vies com podria ser estar 2-3 vegades per sobre de la desviació estàndard) i un cop eliminades les puntuacions extremes, es passa a fer un anàlisi estadístic, no només de la àrea crítica (és a dir, el centre del nostre estudi) i l'àrea *spillover* (és a dir, l'àrea immediatament després de la crítica), sinó també de les tasques distractoras. En *event related potentials* (o ERPs) el participant llegeix en una pantalla una frase (presentada paraula per paraula) mentre porta un casquet amb electròdes que mesuren l'activitat elèctrica cerebral i s'analitzen els patrons generats en llegir els estímuls. El primer que han de fer els experimentadors es decidir de quants elèctrodes enregistraran l'activitat i de quants d'aquests utilitzaran la informació obtinguda, variant normalment entre 3 i al voltant de 9 (Tanner, Osterhout & Herschensohn, 2009). Abans de fer l'anàlisi estadístic també s'ha de reduir el soroll produït durant l'enregistrament, ja que un únic elèctrode obté informació d'un número molt alt de neurones i moviments musculars involuntaris com pot ser el parpelleig, donant ones d'activitat no rellevants per l'estudi. El més habitual és que es comparin amplituds d'ona en diferents elèctrodes, per condicions experimentals i entre participants en els períodes de temps que siguin rellevants per a l'estudi. Alguns d'aquests períodes de temps són de 300 a 500 ms després de l'estímul per a un N400 i 500 a 800 ms després de l'estímul experimental per un P600 (Tanner et al., 2009).

Per últim, el eye-tracking es considera un mètode molt natural i la seva resolució temporal molt alta fa que sigui considerada una de les mesures més sensibles del processament moment per moment disponibles actualment. En aquest es presenten frases senceres en una pantalla i, mentre el participant llegeix, una càmera segueix els moviments dels ulls. Al contrari del que es pensa comunament, el moviment dels ulls en llegir una frase no es fa en una línia suau

i recta d'esquerra a dreta, sinó que és molt més turbulent. Els ulls es mouen en una sèrie de salts, anomenats *saccades*, separats per períodes curts en els que els ulls es queden relativament quiets, que són les fixacions. Les *saccades* són moviments ràpids (fan amb una durada d'entre 20 i 40 ms) i normalment mou els ulls de set a nou caràcters (tot i que poden ser d'un únic caràcter o tan llargues com vint) cap a l'esquerra o la dreta. Entre el 85 i el 90% de les *saccades* mouen els ulls cap endavant en el text (és a dir, cap a la dreta) mentre que el 10-15% restant els mouen cap enrere. Aquestes segones s'anomenen regressions i solen ser el resultat de dificultats de comprensió o errors en la programació de *saccades* (quan aquestes fan que els ulls es saltin una regió que és d'interès). En comparació amb les *saccades*, les fixacions són més llargues, tenint una durada habitual d'entre 200 i 250 ms i, a diferència de les *saccades*, els lectors extreuen informació significativa sobre el text durant les fixacions. Els temps de fixació i les regressions són les dues fonts de dades que els investigadors utilitzen més sovint per fer inferències sobre els processos cognitius involucrats en la lectura. D'aquestes dues fonts, es sol donar més atenció a les fixacions ja que constitueixen el 90% dels temps de lectura (essent el 10% restant *saccades*). D'acord amb la teoria *eye-mind* (Just & Carpenter, 1980), es creu que el temps que triga una persona en llegir una paraula reflecteix el temps necessari per processar-la i, per tant, les paraules més difícils de processar resultaran en temps de lectura més alts mentre que passa el contrari amb les paraules més fàcils de processar. Paraules més complexes de processar també poden donar lloc a regressions, pel qual aquestes solen ser un focus secundari en les anàlisi.

En les frases que es presenten durant l'experiment, sempre hi ha una regió crítica, que seria la regió principal d'interès i la que l'investigador manipula per produir efectes. Tota la resta són regions no crítiques.

A l'hora de recopilar i analitzar els resultats, els investigadors tenen diferents mesures separades entre *early* i *late*. Es creu que les mesures *early* reflecteixen les anàlisis i interpretacions inicials de cada paraula quan el lector se la troba per primer cop. Algunes d'aquestes mesures serien: la *first fixation duration* és la durada de la primera fixació que fa un lector en una paraula; la *gaze duration* és la suma de les durades de totes les fixacions en una paraula fins que es surt d'aquesta zona ja sigui per l'esquerra o la dreta; el *first-pass time* és el nom que se li dona a la *gaze duration* quan la regió crítica consisteix de més d'una

paraula. D'altra banda es creu que les mesures *late* ens indiquen un re-anàlisi degut a la dificultat del processament. Un exemple d'aquest segon tipus de mesura seria el *total time* que es refereix a la suma de les durades de totes les fixacions fetes en una paraula, incloent el temps dedicat a rellegir la mateixa. L'investigador decidirà quines mesures triar i si les regressions dins o fora d'una regió li seran informatives, però la majoria escullen un parell de mesures early i almenys una mesura late.

Per a cada mesura i regió es perdran alguns valors. Quan falten valors en les mesures early o hi ha *track loss*, es deixen com a desapareguts en lloc de substituir-los per altres valors. Quan aquesta pèrdua de valors és per a mesures late, es sol substituir aquest per zero, reflectint així que els lectors no van necessitar rellegir una regió prèviament fixada. També és comú eliminar les dades d'aquells assajos en que els participants han respost incorrectament a la pregunta post-estímul.

Per últim, també es retallen els valors extrems que podrien afectar els resultats de l'estudi. No hi ha un valor màxim i mínim de fixació concret que ens digui a on s'ha de fer un tall sinó que aquest varia d'un estudi a un altre, però normalment oscil·la entre menys de 50 a 100 ms per les fixacions curtes i superior a 800 ms aproximadament. El fet de tallar fixacions molt llargues i considerar-les com a valors erronis es deu també a que aquestes poden estar reflectint processos diferents al que s'està estudiant com la fatiga o distracció.

Estudis previs sobre la concordança/no concordança de gènere:

En estudis electrofisiològics com el de Aleman Banon & Rothman (2016) es pot observar com els participants són capaços de distingir entre oracions gramaticals i oracions amb no concordança de gènere independentment de la tasca en espanyol (fet que es pot concloure perquè les puntuacions dels participants en tasques de judici gramatical van ser molt altes). Això ho veiem també en l'estudi de Molnar et al. (2020), on demostren que infants de 18 mesos d'edat (monolingües i bilingües que aprenen espanyol) són capaços de discernir entre seqüències d'article-nom amb concordança de gènere o sense concordança. En aquest estudi van poder observar més concretament que els nadons monolingües espanyols semblaven preferir les frases correctes, mentre que els bilingües tenien una preferència per les incorrectes. Aquesta preferència es va determinar pel temps que escoltaven les frases exposades (quant més temps escoltant, més preferència)(Molnar et al., 2020). En un altre estudi amb

bilingües de basc i espanyol (on un grup dels participants tenien el basc com la seva llengua dominant mentre que un altre grup tenia l'espanyol) Caffarra et al. (2015) van poder observar que el domini lingüístic en una llengua juga un paper important en determinar el pes de la via lèxica que utilitzen en el còmput de la concordança. Els infants que tenien l'espanyol com a llengua dominant, acostumats a parlar espanyol diàriament, poden confiar en una representació lèxica plenament arrelada en el gènere espanyol, sense la necessitat de calcular el gènere seguint senyals formals disponibles a una frase. D'altra banda, en els infants que tenien el basc com a llengua dominant, els senyals formals influeixen més durant el còmput de concordança perquè la representació lèxica del gènere és menys estable.

Aquest resultat és congruent amb resultats d'estudis conductuals anteriors que mostraven rendiments asimètrics entre noms transparents i opacs amb bilingües altament competents (Bordag et al., 2006). A més a més, no només dona importància a la simple presència d'una segona llengua, sinó que assenjala també la rellevància del domini lingüístic i l'ús de la llengua (Caffarra et al., 2015).

Hi ha un gran nombre d'estudis amb ERPs, tant en espanyol com en altres idiomes que han demostrat que, quan hi ha una violació de la concordança de gènere, es dona un P600, sovint robust, entre 500 ms i 1000 ms. Aquest mateix resultat el veiem en un estudi de marcatge de gènere morfològic també, en el que es va poder observar el P600. Aquest es va veure afectat de forma diferent quan les violacions de gènere eren en femení (el que significaria que té un marcadore semàntic), condició per la qual el P600 sorgeix abans que per la condició masculina (és a dir, adjectius que no tenen un marcadore) (Aleman Banon & Rothman, 2016).

Un estudi que examinava el rol de la semàntica i els processos de reparació en no concordança de gènere entre article i nom en Italià, també va demostrar que un augment en la complexitat de reparació porta a un augment també de l'amplitud del P600 en les fases més tardanes d'aquest (Popov et al., 2020). En anteriors estudis sobre monolingües espanyols (Caffarra & Barber, 2015; Caffarra, et al., 2015) s'ha demostrat l'efecte de la transparència i la concordança de gènere, obtenint resultats que demostraven que els noms transparents van provocar una negativitat més gran que els noms opacs al voltant dels 400 ms (això és, modulacions al N400). En aquest estudi també van

veure un patró LAN-P600 per a les no concordances de gènere tant en noms transparents com opacs el que ens indica que, per els monolingües, la terminació semàntica del noms no té un gran impacte en el còmput de concordança (Caffarra & Barber, 2015; Caffarra, et al., 2015). En un estudi també de transparència i concordança però aquest cop amb infants bilingües (amb basc o espanyol com llengua dominant) van veure que els errors en noms transparents van provocar un patró LAN-P600 més gran, mentre que les violacions de gènere dels noms opacs només van tenir un efecte P600. Els bilingües que tenien l'espanyol com a llengua dominant van obtenir resultats similars als dels monolingües espanyols (Caffarra & Barber, 2015), mostrant respostes LAN-P600 per la no concordança de gènere tant en noms transparents com en opacs. A més a més, els dos grups bilingües (el grup que tenia l'espanyol com a llengua dominant i el grup que tenia el basc com a llengua dominant) van mostrar un efecte de transparència entre els 350 ms i els 500 ms, el que ens indica que eren sensibles a la presència dels senyals de gènere disponibles (Caffarra et al., 2015). En estudis de concordança de gènere en holandès veiem resultats similars, on el P600 es provocat per la manipulació de no concordances de gènere entre article-nom. En aquest cas, però, no es va observar l'efecte del LAN, fet que concorda amb molts altres estudis i van en contra de molts altres (Popov & Bastiaanse, 2018).

Hem vist, per tant, que els resultats més freqüents en estudis amb ERPs de no concordança de gènere son els efectes LAN i P600. Beatty-Martínez et al. (2020) van dur a terme un estudi per analitzar les diferències particulars entre les reaccions a no concordances amb noms masculins i femenins en espanyol. En aquest, van trobar l'efecte LAN i P600 per les violacions de gènere masculines però només un efecte P600 per a les del gènere femení, així com una distribució més amplia d'aquest últim efecte en les condicions de concordança per als noms femenins que per als masculins.

Veient aquests resultats i una resposta bifàsica (es a dir, efectes LAN i P600 de mida relativament igual) només per la condició de no concordança masculina, van fer un anàlisi de les diferències individuals, seleccionant els elèctrodes més representatiu de cada efecte (el F3 per el LAN i el Pz per el P600).

Van trobar que els participants que mostraven un P600 predominant, mostraven un LAN menys robust, i viceversa, però també l'associació contrària al comparar els dos efectes entre condicions. Les amplituds mitges de LAN estaven

inversament correlacionades en les dues condicions de gènere i aquest efecte per als noms masculins va mostrar una correlació positiva amb l'efecte P600 per als noms femenins. Això el que ens voldria dir és que els participants que van tenir un efecte LAN predominant per a la condició de no concordança en noms masculins, van tenir una P600 de igual forma predominant per a la condició de no concordança en noms femenins.

En mesures individuals en que avaluaven si existeix alguna diferència en el rang o tipus de variabilitat provocada per les condicions de l'oposició, van veure que no va haver-hi més sensibilitat a les violacions de gènere entre noms masculins i femenins, i van calcular l'Índex de Dominància de Resposta o *Response Dominance Index* (RDI). El RDI és una mesura de la preponderància relativa dels participants de una negativitat sobre una positivitat, independentment de la magnitud, de la qual veiem els resultats en la Figura 7.

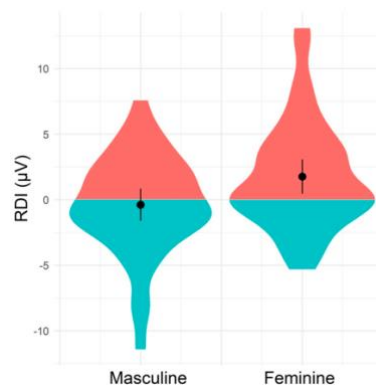


Figura 7. Representació gràfica de la distribució dels valors de RDI en funció del gènere del nom.

Les puntuacions més negatives indicarien una superioritat de l'efecte LAN i, les més positives, de l'efecte P600, mentre que els valors propers a zero serien una resposta bifàsica. Com es pot observar en la Figura 7, les puntuacions tant per la condició de violació de gènere masculina com per la femenina es troben al voltant de zero, amb la diferència que els valors de RDI per als noms masculins s'allarguen en direcció negativa i els valors per noms femenins ho fan en direcció positiva. Els valors RDI també els va permetre recolzar el fet que hi ha una relació asimètrica en la codificació morfològica del gènere (tenint els noms femenins i masculins efectes diferents en el processament) i els efectes P600 relacionats amb la descomposició morfològica i el re-anàlisi, en obtenir valors de RDI significativament més positius per als noms femenins que per als masculins.

2. Objectius

Un dels objectius principals d'aquest estudi és investigar l'efecte que té la no concordança de gènere semàntic en el processament de frases en català.

De la mateixa forma, volem veure les (possibles) diferències quan la no concordança de gènere es dona amb un nom propi o comú.

També volem contribuir a la literatura sobre la concordança i el marcatge de gènere a través de l'anàlisi de moviments oculars.

3. Metodologia

Participants

Van participar en l'experiment trenta quatre nadius catalans (27 dones i 7 homes), d'entre 20 i 35 anys (amb una mitja d'edat de 22.3 i una desviació típica de 3.2). Eren alumnes universitaris i van participar a l'experiment per un crèdit extra a una assignatura. Tenien visió normal o bé corregida per a que sigui normal (és a dir, cap participant tenia dèficits visuals). Tots van donar el seu consentiment informat abans de dur a terme l'experiment i desconeixien els objectius de l'experiment.

Materials

En crear materials per a un experiment d'eye-tracking s'han de tenir en compte diverses consideracions. Per assegurar-se que els efectes que es veuen en analitzar els resultats venen donats per els efectes que estan essent estudiats i no efectes externes, el que es fa es crear la mateixa frase en diferents versions o condicions per a cada nivell de la variables (o variables) independent. Per tant, les frases seran iguals lèxicament (tindran les mateixes paraules) i l'única diferència entre elles serà l'aspecte a estudiar, com en l'exemple següent:

- (1) a. Jo crec que el nen va guanyar un concurs de pintura.
- b. Jo crec que la nen va guanyar un concurs de pintura.

A aquest conjunt de frases se l'anomena *item* i cada item apareix en dues versions, en aquest cas, una amb concordança de gènere (1a) i una amb no concordança de gènere (1b). Aquest seria un exemple per a un experiment que estudia una única variable independent de dos nivells, però si tinguéssim dues variables independents amb dos nivells per cada una d'elles, cada item tindria 4 frases diferents. En l'exemple veiem com són les frases utilitzades en el nostre experiment, essent dues les variables independents. La primera variable independent és tipus de nom (nom propi vs. nom comú), mentre que la segona és concordança de gènere (concordança vs. no concordança).

- (2) a. Jo crec que el nen va guanyar un concurs de pintura (nom comú, concordança).
b. Jo crec que la nen va guanyar un concurs de pintura (nom comú, no concordança).
c. Jo crec que el Marc va guanyar un concurs de pintura (nom propi, concordança).
d. Jo crec que la Marc va guanyar un concurs de pintura (nom propi, no concordança).

Serà important també que les versions dintre de l'ítem siguin el més semblants possibles entre elles, és a dir, que hi hagi consistència. El processament de la regió crítica (la regió que manipulem per estudiar la variable, en el nostre cas, la regió del nom) acostuma a continuar en les paraules immediatament després, fet que es coneix com l'efecte de *spillover*, i per tant, fent que hi hagi consistència dintre dels ítems, ens assegurem un cop més que el resultat ve per les variables que estem estudiant i no factors externs.

Un altre factor a tenir en compte és el nombre d'ítems que incloure en l'experiment. Una de les consideracions per decidir-ho és que es perden algunes dades durant l'experiment (ja sigui per errors o distracció del participant), normalment entre un 8 i 15%. A més a més, el mateix participant mai llegirà més d'una versió del mateix ítem (és a dir, si llegeix (1a) no llegirà (1a)) ja que llegir més d'una podria causar l'efecte de repetició i fer que els participants no responguin de forma natural perquè ja han vist l'ítem anteriorment. Per minimitzar aquest efecte, es divideixen les versions dels ítems en diferents llistes de presentació, fent que un participant llegeixi només les frases d'una d'aquestes llistes. El nombre de llistes de presentació és el mateix que el nombre de versions dintre d'un mateix ítem, i per tant, en un exemple com el de (2) les llistes serien les següents:

Llista 1: 1a, 2b, 3c, 4d, 5a, 6b, 7c, 8d, etc.

Llista 2: 1b, 2c, 3d, 4a, 5b, 6c, 7d, 8a, etc.

Llista 3: 1c, 2d, 3a, 4b, 5c, 6d, 7a, 8b, etc.

Llista 4: 1d, 2a, 3b, 4c, 5d, 6a, 7b, 8c, etc.

Per tant, per obtenir dades d'almenys entre 6 a 10 ítems tenint en compte les que es perdran, en un estudi com el nostre es crearan entre 32 i 48 ítems. El posicionament de cada ítem dintre de la llista de presentació ha d'estar balancejada, de forma que cap ítem o versió aparegui de forma consistent en

aproximadament la mateixa posició durant tot l'experiment. El llistat complet de materials experimentals del nostre estudi es pot trobar a l'apèndix.

A més a més dels items experimentals, també se li presentaran als participants frases distractores o fillers, això és, frases que no tenen res a veure amb l'objectiu de l'experiment i que tenen per funció evitar que el participant descobreixi la manipulació experimental. Després d'algunes de les frases es presentarà una pregunta de comprensió per assegurar-nos que els participants llegeixen les frases prestant atenció. Les preguntes de comprensió estan sempre relacionades amb el contingut de la oració (com es veu en l'exemple de (3), tot i que també és comú utilitzar preguntes de judici d'acceptabilitat).

(3) L'Alba ha d'agafar dos autobusos per anar a classe (oració).

Va l'Alba amb metro a classe? (pregunta).

Tal i com ja hem dit, els *fillers* són frases que no tenen relació amb el que s'està estudiant i no busquen provocar cap tipus d'efecte de processament (com seria presentar només l'oració de (3)). Tot i que de vegades el terme filler i distractor s'utilitzen indistintament, es poden utilitzar els dos per diferenciar entre diferents tipus d'item no crític. L'objectiu d'incloure fillers és amagar dels participants els items crítics i, per tant, els objectius específics del nostre estudi i minimitzar els efectes de tasca (és a dir, efecte de repetició o que els participants adoptin una estratègia de processament no natural).

Un últim punt a tenir en compte és que per minimitzar l'efecte de familiaritat de la tasca, abans de començar la sessió, també es presentaran entre 5 i 10 items de pràctica, sense relació amb els estímuls crítics. Per minimitzar també l'efecte de fatiga, es pot limitar el número d'assajos administrats en una sessió, fent que, entre estímuls experimentals i fillers o distractors, les frases llegides no superin un total de 120 a 160 (sense contar amb els items de pràctica) i exigint als participants que facin una o més pauses de lectura durant l'experiment.

Procediment

Els participants van passar l'experiment individualment en una habitació insonoritzada. Es va utilitzar un eye-tracker, concretament un EyeLink 1000 (SR Research), per enregistrar els moviments oculars mentre els participants llegien les frases. Es van presentar els estímuls a una distancia constant de 60 cm de una pantalla d'ordinador de 19 polzades, establerta a una resolució de 1024 x 768 píxels. El visionat era binocular, però només s'enregistrava contínuament els moviments de l'ull dret a una freqüència de mostreig de 1000Hz. Les frases es

van presentar centrades, en una sola línia i en minúscules negres (Arial, 24) implementades per el software SR Research Experiment Builder. Abans d'iniciar l'experiment, els participants van llegir i signar un formulari de consentiment informat, després van llegir les instruccions i també van completar una breu pràctica, que consistia en sis frases, per a familiaritzar-se amb el procediment. Es va dur a terme la calibració abans de l'experiment (abans i després de la fase de pràctica) i a la meitat del mateix després d'una breu pausa. L'experiment va tenir una durada aproximada de 30 minuts.

Anàlisi de les dades

Els assajos amb *track loss* o error en la recollida de dades no es van incloure a l'anàlisi. A més a més, es van eliminar les fixacions més curtes que 80 ms i més llargues que 1200 ms. Les dades dels moviments oculars es van analitzar per la regió crítica i post-crítica. La regió crítica comprenia el nom (propi o comú) i la regió post-crítica les dos paraules que el seguien.

Es van computar quatre mesures de seguiment ocular en la regió crítica. La *first fixation duration* és la duració de la primera fixació en una regió d'interès en la primera passada durant la lectura. *Gaze duration* és la suma dels temps de fixació abans que els ulls surtin de la regió, ja sigui cap a l'esquerra o la dreta. *Go-past time* és la suma de la duració de les fixacions fins a que hi ha una sortida cap a la dreta, incloent qualsevol fixació a la esquerra de la regió crítica com a resultat de moviments oculars regressius. El total *reading time* és la suma de totes les fixacions, incloent la primera, segona i les següents "passes". Es van analitzar les mateixes mesures de moviments oculars en la regió post-crítica, exceptuant la *first fixation duration*, que es va excloure degut a que aquesta regió abasta més d'una paraula.

4. Resultats

La precisió global a les preguntes de comprensió va ser del 90,1% (SD = 4,8, rang = 100-80), indicant que els participants paraven atenció al contingut de les frases. Donat que els índex de precisió van ser superiors al 75% en tots els participants, es va incloure a tots els participants en l'anàlisi.

Regió crítica

La Taula 1 proporciona informació sobre la mitjana (i la SD) dels temps de lectura de les quatre mesures a la regió crítica per a les quatre condicions experimentals.

Nom	Concordança	FFD	GD	RPD	TOTAL
Comú	Conc	209,4 (37,6)	324,7 (111,5)	410,0 (172,2)	506,9 (196,3)
	NoConc	226,4 (44,8)	365,1 (156,5)	647,2 (252,7)	765,5 (238,9)
Propi	Conc	170,0 (34,7)	196,9 (51,6)	247,1 (87,5)	285,2 (94,4)
	NoConc	184,9 (43,9)	242,0 (103,6)	401,6 (172,8)	439,8 (137,5)

Taula 3. Mitjana dels temps de lectura (en ms) i desviació estàndard a la regió crítica.

Es va observar un efecte principal de Tipus de Nom a first fixation duration ($F(1,33)=35,75$, $p<.001$, $\eta^2=.52$), gaze duration ($F(1,33)=51,61$, $p<.001$, $\eta^2=.61$), regression path duration ($F(1,33)=45,87$, $p<.001$, $\eta^2=.58$) i total reading time ($F(1,33)=82,15$, $p<.001$, $\eta^2=.71$). Aquest efecte principal es deu a que els temps de lectura són superiors pels noms comuns que pels noms propis. Aquest efecte pot explicar-se pel fet que els noms comuns són menys freqüents que els noms propis ($t(78)=-3,39$, $p=.001$) i són més llargs, en número de lletres, que els noms propis ($t(78)=9,52$, $p<.001$).

També es va observar un efecte principal de Concordança a first fixation duration ($F(1,33)=9,67$, $p=.004$, $\eta^2=.22$), gaze duration ($F(1,33)=10,38$, $p=.003$, $\eta^2=.23$), regression path duration ($F(1,33)=67,90$, $p<.001$, $\eta^2=.67$) i total reading time ($F(1,33)=73,99$, $p<.001$, $\eta^2=.69$). Aquest efecte principal es deu a que els temps de lectura en la condició no-concordança són significativament superiors a la condició de concordança de gènere.

Per últim, es va observar que la interacció entre Tipus de Nom i Concordança de Gènere no era significativa a first fixation duration i a gaze duration ($Fs<1$), però sí era significativa a regression path duration ($F(1,33)=6,36$, $p=.017$, $\eta^2=.16$) i a total reading time ($F(1,33)=5,52$, $p=.025$, $\eta^2=.14$). Les comparacions per parelles (corregides pel mètode Bonferroni) van mostrar que l'efecte de Concordança era més gran als noms comuns (a RPD, diferència de 237 ms, $p<.001$, $\eta^2=.69$, i a TOTAL, diferència de 249 ms, $p<.001$, $\eta^2=.57$) que als noms propis (a RPD, diferència de 154 ms, $p<.001$, $\eta^2=.44$, i a TOTAL, diferència de 154 ms, $p<.001$, $\eta^2=.52$).

Regió Post-crítica

La Taula 2 proporciona informació sobre la mitjana (i SD) dels temps de lectura de les quatre mesures a la regió post-crítica per a les quatre condicions experimentals.

Nom	Concordança	GD	RPD	TOTAL
Comú	Conc	418,3 (115,9)	493,6 (145,6)	590,2 (190,4)
	NoConc	396,4 (137,5)	721,9 (274,1)	709,1 (235,5)
Propi	Conc	436,6 (127,3)	484,7 (159,3)	594,4 (181,7)
	NoConc	400,9 (146,2)	657,9 (221,8)	596,9 (180,3)

Taula 4. Mitjana dels temps de lectura (en ms) i desviació estàndard a la regió post-crítica

L'efecte principal de Tipus de Nom no va ser significatiu a gaze duration ($F < 1$), va ser marginalment significatiu a regression path duration ($F(1,33) = 3,38$, $p = .075$, $\eta^2 = .09$), i va ser significatiu a total reading time ($F(1,33) = 6,50$, $p = .016$, $\eta^2 = .16$). Aquest efecte principal es deu a que els temps de lectura dels noms propis són menors als temps de lectura dels noms comuns. L'efecte principal de Concordança de gènere va ser significatiu a gaze duration ($F(1,33) = 6,99$, $p = .012$, $\eta^2 = .17$), a regression path duration ($F(1,33) = 55,27$, $p < .001$, $\eta^2 = .62$), i a total reading time ($F(1,33) = 10,90$, $p = .002$, $\eta^2 = .24$). Aquest efecte principal es deu a que la condició de concordança generava temps de lectura significativament menors que la condició de no-concordança. Per últim, la interacció entre Tipus de Nom i Concordança de Gènere no va ser significativa ni a gaze duration ni a regression path duration ($F_s < 1$). La interacció sí va ser significativa a la mesura total reading time ($F(1,33) = 5,52$, $p = .025$, $\eta^2 = .14$). Les comparacions per parells van assenyalar que en el cas dels noms propis no hi havia diferències entre les condicions de concordança i no concordança ($p = .93$), mentre que en el cas dels noms comuns les diferències eren significatives ($p = .001$), observant-se temps de lectura significativament menors en la condició de concordança que en la condició de no concordança.

5. Discussió i conclusions

Vam dur a terme un experiment amb eye-tracker per investigar l'efecte de la no concordança de gènere en el processament de frases en català i veure possibles diferències entre noms propis i noms comuns. Pel que fa als noms propis, només vam contemplar noms propis que solament tenen una forma, sigui en masculí o femení, evitant els noms propis que poden tenir tant una versió en masculí (e.g., Joan) com en femení (e.g., Joana). Respecte als noms comuns, vam buscar noms que admeten dues formes, una pel femení (e.g., nena) i una altra pel masculí (e.g., nen). La idea principal de l'estudi era veure si les vies de reparació de la no concordança entre un article i el nom eren diferents entre els noms

propis que només admeten una forma (sigui masculina o femenina) i els noms comuns que poden realitzar-se tant en gènere masculí com en femení. En el cas dels noms propis amb una única forma (e.g., *Sergi*, en el cas de gènere masculí, o *Núria*, en el cas de gènere femení), s'esperava que les dificultats de processament fossin menors, ja que la única via de reparació de l'error gramatical era modificar el gènere de l'article, que no pas en el cas dels noms comuns (e.g., *el nena/*la nen) on hi ha dues opcions per reparar l'error gramatical. Una primera opció implicaria modificar el gènere de l'article, mentre que la segona opció implicaria modificar el gènere del nom. Vam hipotetitzar que el fet de tenir dues possibles vies de reparació, en contraposició a disposar d'una única via de reparació, implica que el sistema de processament té més incertesa i això es pot manifestar en un increment en els temps de lectura.

El que mostren els resultats és que tant a la regió crítica com a la post-crítica hi ha un efecte principal de la variable tipus de nom. Aquest efecte principal indica que els temps de lectura eren superiors en la condició de noms comuns que en la condició de noms propis.. Aquest efecte principal es pot explicar pel fet que els noms propis eren més freqüents i eren més curts que els noms comuns. També vam observar un efecte principal de la variable concordança. Els temps de fixació són més grans en les frases en què no hi ha concordança de gènere que en les que sí que hi ha concordança de gènere entre article i nom. A més, vam poder observar que l'efecte de concordança era més gran als noms comuns que als noms propis (en regression path duration i total reading time) en la regió crítica, mentre que en la regió post-crítica, no va haver-hi diferències significatives entre les condicions de concordança i no concordança als noms propis però sí en el cas dels noms comuns (amb temps de lectura menors en la condició de concordança que en la condició de no-concordança).

L'efecte de les oracions amb no concordança de gènere ja s'havia confirmat amb una quantitat d'estudis anteriors, molts d'ells amb ERPs, en els que les condicions de no-concordança generaven un efecte P600, relacionat amb la descomposició morfològica i el re-anàlisi. Aquest efecte està present en quasi tots (si no tots) els estudis, reforçant la teoria dels processos que ha de fer el parser per reparar una frase quan es troba amb una violació de la concordança. D'altra banda, alguns estudis troben també l'efecte LAN, tot i que no és tant regular com l'anterior.

Consideracions:

La tendència en experiments de processaments de frases és tenir un mínim de 12 participants per condició, el que donaria 48 participants en experiments amb 4 nivells de cada item. Per tant, en aquest experiment el nombre de participants no arriba a aquest mínim, fet que s'haurà de tenir en compte a l'hora de valorar els resultats i considerar augmentar la mostra per a futurs estudis.

6. Bibliografia

- Alemán Bañón, J.A., Rothman, J. (2016). The role of morphological markedness in the processing of number and gender agreement in Spanish: an event-related potential investigation. *Language Cognition and Neuroscience*, 31 (10), 1273-1298. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000387912600004>
- Beatty-Martínez, A.L., Bruni, M.R, Bajo, M.T, Dussias, P.E. (2020). Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers. *Psychophysiology*, 58 (8).
- Bordag, D., Opitz, A., Pechmann, T. (2006). Gender processing in first and second languages. The role of noun termination. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32, 1090-1101. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000240160800010>
- Caffarra, S., Barber, H.A. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, 1605, 83-92. <https://www.sciencedirect.com/science/article/pii/S0006899315001183>
- Caffarra, S., Barber, H., Molinaro, N., Carreiras, M. (2017). When the end matters: influence of gender cues during agreement computation in bilinguals. *Language Cognition and Neuroscience*, 32 (9), 1069-1085. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000411462800001>
- Caffarra, S., Siyanova-Chanturia, A., Pesciarelli F., Vespignani, F., Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, 52, 1019-1030. <https://onlinelibrary.wiley.com/doi/10.1111/psyp.12429>
- Fernández, E.M., Cairns, H.S. (2010). *Fundamentals of psycholinguistics*. John Wiley & Sons.

- Just, M.A., Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87 (4), 329-354. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:A1980KB51900001>
- Keating, G.D. (2013). Eye-tracking with text. In *Research methods in second language psycholinguistics* (pp.69-92). Routledge.
- Keating, G.D., Jegerski, J. (2015). Experimental Designs in Sentence Processing Research: A methodological Review and User's Guide. *Studies in Second Language Acquisition*, 37 (1), 1-32. doi:10.1017/S0272263114000187
- Molnar, M., Bañón, J.A., Mancini, S., Caffarra, S. (2021). The processing of Spanish Article-Noun Gender Agreement by Monolingual and Bilingual Toddlers. *Language and Speech*, 64 (4), 980-990. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000600154200001>
- Popov, S., & Bastiaanse, R. (2018). Processes underpinning gender and number disagreement in Dutch: An ERP study. *Journal of Neurolinguistics*, 46, 109-121. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000427219200008>
- Popov, S., Miceli, G., Curcic-Blake, B., Bastiaanse, R. (2020). The role of semantics and repair processes in article-noun gender disagreement in Italian: An ERP study. *Brain and Language*, 206. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000537623500002>
- Tanner, D., Osterhout, L., Herschensohn, J. (2009). Snapshots of grammaticalization: Differential electrophysiological responses to grammatical anomalies with increasing L2 exposure. *Proceedings of the 33rd annual Boston University conference on language development*, 1 i 2, 528-539. <https://www-webofscience-com.sabidi.urv.cat/wos/alldb/full-record/WOS:000266036800045>
- Vigliocco, G., Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, 40, 455-478. <https://www.sciencedirect.com/science/article/pii/S0749596X98926249>

7. Apèndix

Llistat complet dels materials experimentals. Les versions (a) pertanyen a la condició nom comú i concordança; les versions (b) pertanyen a la condició nom comú i no concordança; les versions (c) pertanyen a la condició nom propi i concordança; les versions (d) pertanyen a la condició nom propi i no concordança.

- 1) a. Jo crec que el nen va guanyar un concurs de pintura.
b. Jo crec que la nen va guanyar un concurs de pintura.
c. Jo crec que el Marc va guanyar un concurs de pintura.
d. Jo crec que la Marc va guanyar un concurs de pintura.
- 2) a. Jo recordo que la nena va suspendre un examen a l'escola.
b. Jo recordo que el nena va suspendre un examen a l'escola.
c. Jo recordo que la Mireia va suspendre un examen a l'escola.
d. Jo recordo que el Mireia va suspendre un examen a l'escola.
- 3) a. Nosaltres imaginem que el professor haurà deixat els apunts a copisteria.
b. Nosaltres imaginem que la professor haurà deixat els apunts a copisteria.
c. Nosaltres imaginem que el Joel haurà deixat els apunts a copisteria.
d. Nosaltres imaginem que la Joel haurà deixat els apunts a copisteria.
- 4) a. Tu esperes que la cambrera hagi estat amable amb els clients.
b. Tu esperes que el cambrera hagi estat amable amb els clients.
c. Tu esperes que la Núria hagi estat amable amb els clients.
d. Tu esperes que el Núria hagi estat amable amb els clients.
- 5) a. Vosaltres penseu que el president hauria d'haver parlat amb el gerent.
b. Vosaltres penseu que la president hauria d'haver parlat amb el gerent.
c. Vosaltres penseu que el David hauria d'haver parlat amb el director.
d. Vosaltres penseu que la David hauria d'haver parlat amb el director.
- 6) a. Tu imaginaves que la monitora arribaria a l'hotel abans que vosaltres.
b. Tu imaginaves que el monitora arribaria a l'hotel abans que vosaltres.
c. Tu imaginaves que la Clara arribaria a l'hotel abans que vosaltres.

- d. Tu imaginaves que el Clara arribaria a l'hotel abans que vosaltres.
- 7) a. Nosaltres sabíem que el rector tenia una casa a la platja.
b. Nosaltres sabíem que la rector tenia una casa a la platja.
c. Nosaltres sabíem que el Jaume tenia una casa a la platja.
d. Nosaltres sabíem que la Jaume tenia una casa a la platja.
- 8) a. Jo recordo que la clienta va pagar la consumició en efectiu.
b. Jo recordo que el clienta va pagar la consumició en efectiu.
c. Jo recordo que la Marta va pagar la consumició en efectiu.
d. Jo recordo que el Marta va pagar la consumició en efectiu.
- 9) a. Vosaltres creieu que el director no parla prou amb els treballadors.
b. Vosaltres creieu que la director no parla prou amb els treballadors.
c. Vosaltres creieu que el Sergi no parla prou amb els treballadors.
d. Vosaltres creieu que la Sergi no parla prou amb els treballadors.
- 10) a. Elles afirmen que la professora va ocultar la informació al consell.
b. Elles afirmen que el professora va ocultar la informació al consell.
c. Elles afirmen que la Sònia va ocultar la informació al consell.
d. Elles afirmen que el Sònia va ocultar la informació al consell.
- 11) a. Ells neguen que el tutor es comportés grollerament amb els alumnes.
b. Ells neguen que la tutor es comportés grollerament amb els alumnes.
c. Ells neguen que el Rubén es comportés grollerament amb els alumnes.
d. Ells neguen que la Rubén es comportés grollerament amb els alumnes.
- 12) a. Nosaltresensem que la presidenta hauria d'haver dimitit del seu càrrec.
b. Nosaltresensem que el presidenta hauria d'haver dimitit del seu càrrec.
c. Nosaltresensem que la Natàlia hauria d'haver dimitit del seu càrrec.
d. Nosaltresensem que el Natàlia hauria d'haver dimitit del seu càrrec.
- 13) a. Tu volies que el porter deixés les cartes a la bústia.
b. Tu volies que la porter deixés les cartes a la bústia.
c. Tu volies que el Pere deixés les cartes a la bústia.
d. Tu volies que la Pere deixés les cartes a la bústia.

- 14) a. Jo desitjo que la refugiada pugui tornar aviat al seu país.
b. Jo desitjo que el refugiada pugui tornar aviat al seu país.
c. Jo desitjo que la Marina pugui tornar aviat al seu país.
d. Jo desitjo que el Marina pugui tornar aviat al seu país.
- 15) a. Elles volen que el delegat sigui la persona més votada demà.
b. Elles volen que la delegat sigui la persona més votada demà.
c. Elles volen que el Xavi sigui la persona més votada demà.
d. Elles volen que la Xavi sigui la persona més votada demà.
- 16) a. Nosaltres pensàvem que la propietària deixava el cotxe sempre al pàrquing.
b. Nosaltres pensàvem que el propietària deixava el cotxe sempre al pàrquing.
c. Nosaltres pensàvem que la Laia deixava el cotxe sempre al pàrquing.
d. Nosaltres pensàvem que el Laia deixava el cotxe sempre al pàrquing.
- 17) a. Vosaltres dieu que el cuiner té una trajectòria laboral molt bona.
b. Vosaltres dieu que la cuiner té una trajectòria laboral molt bona.
c. Vosaltres dieu que el Lluç té una trajectòria laboral molt bona.
d. Vosaltres dieu que la Lluç té una trajectòria laboral molt bona.
- 18) a. Tu creies que la psicòloga vivia a la plaça del poble.
b. Tu creies que el psicòloga vivia a la plaça del poble.
c. Tu creies que la Blanca vivia a la plaça del poble.
d. Tu creies que el Blanca vivia a la plaça del poble.
- 19) a. Ells esperen que el perruquer trobi aviat un pis de lloguer.
b. Ells esperen que la perruquer trobi aviat un pis de lloguer.
c. Ells esperen que el Boris trobi aviat un pis de lloguer.
d. Ells esperen que la Boris trobi aviat un pis de lloguer.
- 20) a. Elles recorden que la peixatera sempre era la que obria primer.
b. Elles recorden que el peixatera sempre era la que obria primer.
c. Elles recorden que la Rosa sempre era la que obria primer.
d. Elles recorden que el Rosa sempre era la que obria primer.

- 21) a. Nosaltres recordem que el carnisser sempre ens donava unes llaminadures boníssimes.
- b. Nosaltres recordem que la carnisser sempre ens donava unes llaminadures boníssimes.
- c. Nosaltres recordem que el Pol sempre ens donava unes llaminadures boníssimes.
- d. Nosaltres recordem que la Pol sempre ens donava unes llaminadures boníssimes.
- 22) a. Vosaltres voleu que la pintora es presenti al concurs de cuina.
- b. Vosaltres voleu que el pintora es presenti al concurs de cuina.
- c. Vosaltres voleu que la Tània es presenti al concurs de cuina.
- d. Vosaltres voleu que el Tània es presenti al concurs de cuina.
- 23) a. Jo pensava que el pastisser buscava un local més ben situat.
- b. Jo pensava que la pastisser buscava un local més ben situat.
- c. Jo pensava que el Saül buscava un local més ben situat.
- d. Jo pensava que la Saül buscava un local més ben situat.
- 24) a. Tu esperes que la dissenyadora guanyi el primer premi al concurs.
- b. Tu esperes que el dissenyadora guanyi el primer premi al concurs.
- c. Tu esperes que la Raquel guanyi el primer premi al concurs.
- d. Tu esperes que el Raquel guanyi el primer premi al concurs.
- 25) a. Ells creuen que el secretari hauria de convocar aviat una reunió.
- b. Ells creuen que la secretari hauria de convocar aviat una reunió.
- c. Ells creuen que el Ferran hauria de convocar aviat una reunió.
- d. Ells creuen que la Ferran hauria de convocar aviat una reunió.
- 26) a. Elles esperen que la filòsofa doni la conferència inaugural del congrés.
- b. Elles esperen que el filòsofa doni la conferència inaugural del congrés.
- c. Elles esperen que la Susanna doni la conferència inaugural del congrés.
- d. Elles esperen que el Susanna doni la conferència inaugural del congrés.
- 27) a. Jo vull que el dermatòleg obtingui una plaça a l'hospital comarcal.
- b. Jo vull que la dermatòleg obtingui una plaça a l'hospital comarcal.
- c. Jo vull que el Santi obtingui una plaça a l'hospital comarcal.

- d. Jo vull que la Santi obtingui una plaça a l'hospital comarcal.
- 28) a. Nosaltres creiem que la neuròloga es mereix el premi nacional d'investigació.
- b. Nosaltres creiem que el neuròloga es mereix el premi nacional d'investigació.
- c. Nosaltres creiem que la Tamara es mereix el premi nacional d'investigació.
- d. Nosaltres creiem que el Tamara es mereix el premi nacional d'investigació.
- 29) a. Tu penses que el conductor hauria d'haver auxiliat a les víctimes.
- b. Tu penses que la conductor hauria d'haver auxiliat a les víctimes.
- c. Tu penses que el Maurici hauria d'haver auxiliat a les víctimes.
- d. Tu penses que la Maurici hauria d'haver auxiliat a les víctimes.
- 30) a. Elles desitgen que la politòloga presenti aviat la seva tesi doctoral
- b. Elles desitgen que el politòloga presenti aviat la seva tesi doctoral
- c. Elles desitgen que la Sara presenti aviat la seva tesi doctoral
- d. Elles desitgen que el Sara presenti aviat la seva tesi doctoral
- 31) a. Ells pensen que el peixater hauria de tractar millor els clients.
- b. Ells pensen que la peixater hauria de tractar millor els clients.
- c. Ells pensen que el Roger hauria de tractar millor els clients.
- d. Ells pensen que la Roger hauria de tractar millor els clients.
- 32) a. Vosaltres pensàveu que la locutora havia treballat sempre a la ràdio.
- b. Vosaltres pensàveu que el locutora havia treballat sempre a la ràdio.
- c. Vosaltres pensàveu que la Mònica havia treballat sempre a la ràdio.
- d. Vosaltres pensàveu que el Mònica havia treballat sempre a la ràdio.
- 33) a. Nosaltres sabem que el traductor té molts coneixements sobre l'escriptor rus.
- b. Nosaltres sabem que la traductor té molts coneixements sobre l'escriptor rus.
- c. Nosaltres sabem que el Fèlix té molts coneixements sobre l'escriptor rus.
- d. Nosaltres sabem que la Fèlix té molts coneixements sobre l'escriptor rus.
- 34) a. Jo prefereixo que la ballarina vagi a l'hotel de la platja.
- b. Jo prefereixo que el ballarina vagi a l'hotel de la platja.
- c. Jo prefereixo que la Marga vagi a l'hotel de la platja.
- d. Jo prefereixo que el Marga vagi a l'hotel de la platja.

- 35) a. Tu detestes que el regidor aparqui sovint al pas de vianants.
b. Tu detestes que la regidor aparqui sovint al pas de vianants.
c. Tu detestes que el Dídac aparqui sovint al pas de vianants.
d. Tu detestes que la Dídac aparqui sovint al pas de vianants.
- 36) a. Elles imaginem que la vicerectora arribarà abans que comenci el sopar.
b. Elles imaginem que el vicerectora arribarà abans que comenci el sopar.
c. Elles imaginem que la Nerea arribarà abans que comenci el sopar.
d. Elles imaginem que el Nerea arribarà abans que comenci el sopar.
- 37) a. Ells esperen que el coordinador contracti a més personal pel congrés.
b. Ells esperen que la coordinador contracti a més personal pel congrés.
c. Ells esperen que el Jonatan contracti a més personal pel congrés.
d. Ells esperen que la Jonatan contracti a més personal pel congrés.
- 38) a. Vosaltres creieu que la nedadora guanyarà fàcilment les dues competicions internacionals.
b. Vosaltres creieu que el nedadora guanyarà fàcilment les dues competicions internacionals.
c. Vosaltres creieu que la Sofia guanyarà fàcilment les dues competicions internacionals.
d. Vosaltres creieu que el Sofia guanyarà fàcilment les dues competicions internacionals.
- 39) a. Nosaltres dubtem que el funcionari ens deixi fer còpies dels documents.
b. Nosaltres dubtem que la funcionari ens deixi fer còpies dels documents.
c. Nosaltres dubtem que el Fidel ens deixi fer còpies dels documents.
d. Nosaltres dubtem que la Fidel ens deixi fer còpies dels documents.
- 40) a. Vosaltres voleu que la pedagoga presenti els resultats de l'informe final.
b. Vosaltres voleu que el pedagoga presenti els resultats de l'informe final.
c. Vosaltres voleu que la Bàrbara presenti els resultats de l'informe final.
d. Vosaltres voleu que el Bàrbara presenti els resultats de l'informe final.