

# L'ús de les llengües minoritàries en les xarxes socials: el valencià en Twitter

Vanessa Tölke (Freiburg im Breisgau)

La comunicació mitjançant l'ordinador (*Computer Mediated Communication* o CMC) són les interaccions comunicatives realitzades a través d'Internet (McQuail, 2010: 552). L'anàlisi de la llengua emprada en la CMC és un dels nous objectius de la recerca sociolingüística, en la qual també s'inclou l'estudi de les llengües minoritàries, àmbit extremament interessant e innovador. La CMC es realitza fonamentalment a través de xarxes socials com Twitter i Facebook, molt populars entre usuaris de totes les edats encara que són els més joves el sector que n'optimitza l'ús com a eines virtuals de comunicació.

Twitter, representant de la xarxa social per excel·lència, se situa en el centre de la investigació d'aquesta recerca en la qual intentem aproximar-nos a la qüestió de si Twitter, com a mitjà de comunicació, pot contribuir a mantenir una llengua minoritària, com el valencià, i ajudar a normalitzar-ne l'ús.

La utilització del valencià en les xarxes socials encara no ha estat gaire estudiada per part de la comunitat científica; sobretot, cal dir que no s'ha investigat des d'una perspectiva sociolingüística.<sup>1</sup> La situació del valencià és extraordinària perquè, per una banda, s'ha reconegut filològicament i científicament com una varietat del català, però, per una altra banda, existeix una tradició relativament antiga de distingir el valencià del català i, per tant, de considerar el valencià una llengua independent del català (cf. Doppelbauer, 2006: 122ss.). La situació diglòssica valenciana, entre el valencià i el castellà, provoca que la descripció sociolingüística d'aquesta comunitat sigui encara més complexa com a conseqüència de l'aparent doble conflicte politicolingüístic. Conseqüentment, l'estudi de l'ús del valencià en la xarxa

---

1 L'únic cas conegut durant el temps de la recerca és el del comunicador valencià Eugeni Alemany que al seu blog (<<https://ultramarinosalemany.wordpress.com>>) està realitzant una enquesta no oficial sobre l'ús del valencià en les xarxes socials Twitter i Facebook però els resultats encara no s'han publicat (cf. Alemany, 2014).

Twitter que s'ha dut a terme en aquest treball ha suposat un desafiament enorme. Per desenvolupar el tema objecte d'estudi, en aquest treball es segueix la següent estructura.

En primer lloc, es realitza una introducció al món de Twitter i una discussió breu sobre la classificació de la xarxa com a mitjà oral o escrit. En segon lloc es presenta la metodologia amb la qual s'ha investigat l'ús de la llengua en Twitter a la comarca de la Marina Alta (Alacant). En tercer lloc, es presenta l'anàlisi, que està dividida en dues parts. Per un costat, s'ha estudiat la relació entre l'ús del castellà i el valencià a partir d'un corpus basat en converses extretes de Twitter. Per un altre costat, s'han intentat trobar respostes a la qüestió de si els usuaris joves de Twitter s'apropen més en els seus tuits al model de llengua parlada o si més bé fan servir una varietat escrita quan tuitegen (o piulen<sup>2</sup>).

L'objectiu d'aquest treball és oferir una mostra de l'ús del valencià en la xarxa Twitter. En la investigació s'ha tingut en compte la situació diglòssica entre el castellà i el valencià, l'aparent situació conflictiva entre dues varietats de la mateixa llengua, com és el cas del valencià i el català en la Comunitat Valenciana, i el procés de normalització del valencià.

## ■ 1.1 Twitter: Presentació de la xarxa social

La xarxa social Twitter<sup>3</sup> va ser fundada l'any 2006 i té actualment més de 284 milions de comptes actius en tot el món (cf. Kemp, 2014: 4).<sup>4</sup> Des de juliol de l'any 2012, Twitter també és accessible en català (CCMA, 2014).<sup>5</sup> Twitter s'ha definit com una forma de *microblogging* (Java et al., 2009: 118) en què els usuaris poden publicar missatges de text, que hi apareixen cronològicament en forma de blog en els seus perfils amb un límit de 140 caràcters. D'aquesta manera, és possible generar tuits (publicacions) amb temes diversos i de publicar-los i llegir-los en temps real (*op.cit.*: 119). Les publicacions apareixen en el perfil de l'usuari i s'envien automàticament als

---

2 Segons Termcat (2014) els verbs *tuitejar* i *piular* i els seus derivats es poden utilitzar com a sinònims. El mateix passa amb els substantius *tuit* i *piulada*. Per tant, les formes procedents del nom comercial de la xarxa Twitter, com *tuit*, *tuitejar*, *retuit* i *retuitejar*, actuen com a sinònims complementaris de *piulada* i derivats i són emprades en aquest treball.

3 Nom d'origen anglès, derivat de *to tweet*.

4 Altres dades informen que l'any 2010 hi havia 100 milions d'usuaris actius (Crystal, 2011: 54).

5 La xarxa social Facebook també té versió en català igual que altres xarxes com Youtube o Google+.

seguidors, en altres paraules, a altres usuaris que segueixen l'autor del missatge en la xarxa. De manera que és possible que els seguidors reenvien el mateix tuit, és a dir, que facin un retuit (Crystal, 2011: 36–37).

La limitació de caràcters ha provocat que Twitter també hagi estat designat com el sistema SMS d'Internet (Crystal, 2011: 36–37). El límit de 140 caràcters redueix les opcions gramaticals i gràfiques de l'usuari (*op.cit.*: 20). Per aquesta raó en els tuits no es troben estructures sintàctiques molt desenvolupades, sinó més aviat una aproximació a la llengua parlada i informal, regida pel principi d'economia.

## ■ 1.2 La comunicació en Twitter: llengua oral o llengua escrita?

Com que no es obvi si la llengua en Twitter es pot classificar com a llengua parlada o escrita, és necessari un estudi exhaustiu de la llengua usada en aquesta xarxa per poder definir les característiques d'aquesta forma de comunicació. En el context d'aquesta investigació, Twitter s'ha entès com a híbrid entre la comunicació escrita i l'oral en la qual els usuaris imiten conscientment l'oralitat i construeixen així una mena de proximitat fictícia (Kailuweit, 2009: 1).

Malgrat que les publicacions en Twitter a primera vista no semblen dialògiques, aquesta xarxa, igual que el xat (Kailuweit, 2009: 2), pot ser considerada com una nova forma de comunicació virtual basada en l'escriptura, perquè en els tuits es pot establir una interacció natural entre els usuaris amb l'ús de la @ més el nom de l'usuari a l'inici d'un tuit, com podem comprovar en l'exemple (1a) i (1b).

- (1) a. @AB Princesaaaaa per MOOOLTS anys! mos hem de vore fa moooooolt. Muaaa (Parlant DV, tuit 233)  
 b. @DV Piliiii moltíssimes gràcies, si, ja toca voremos que a jp no et vaig vore. T'estime (Parlant AB, tuit 6223)

Els missatges són més curts que en la comunicació oral, no obstant això, s'estableix un intercanvi semblant al de la comunicació oral en què se simula la llengua oral en forma escrita i el receptor pot reaccionar quasi immediatament al tuit rebut. Així, és possible que l'oralitat s'imiti conscientment.

A partir d'aquest fet, sorgeix la pregunta de si es pot fer servir la terminologia de Koch / Oesterreicher (2011) i, per tant, emprar els termes de llengua de la proximitat (*Nähesprache*) i llengua de la distància (*Distanzsprache*).

*che*) i definir la comunicació en Twitter com a oralitat concepacional (*konzeptuelle Mündlichkeit*). Després d'una anàlisi detallada, la comunicació dins Twitter mostra característiques clares de la llengua de la proximitat, com la privacitat, la intimitat i l'espontaneïtat (cf. exemple 1), però s'hi poden trobar també aspectes que, segons el model de Koch / Oesterreicher, es poden classificar en la distància comunicativa com la distància física, l'oficialitat, etc. (Koch / Oesterreicher, 2011: 12).

En relació amb aquesta heterogeneïtat es qüestiona si la terminologia de Koch / Oesterreicher és adequada per a l'àmbit de la comunicació en CMC, especialment en Twitter. En aquest context cal subratllar, tal com defensa Kailuweit (2009: 14) pel que fa al xat, que la majoria dels usuaris volen construir una proximitat fictícia amb les seves contribucions. Per aconseguir-ho, utilitzen instruments que es poden classificar com a escripturalitat medial (*ibid.*).

Com ja s'ha mencionat, es troben també formes comunicatives amb elements de distància en la comunicació de Twitter, com s'observa en l'exemple (2), en què es mostra el compte del diari *Ara* (2015), un espai informatiu que publica tuits sobre notícies actuals.

- (2) El consell de ministres aprova la privatització parcial d'Aena. Un 28% del seu capital s'estrenarà a borsa (ARAeconomia, 23.01.2015)

Aquest tuit no reflecteix la llengua oral i es classificaria com a llengua de la distància. Més endavant veurem que el corpus utilitzat en aquest estudi no correspon a la llengua de la distància perquè es tracta només de perfils d'usuaris joves amb contingut informal. Per tant, la declaració de Kailuweit (2009: 14) sobre la construcció d'una proximitat fictícia sí que es pot confirmar en aquest cas. Així doncs, imitacions de sons com *bebe* o *haha* per a expressar una rialla, l'ús d'acrònims com *lol* o l'expressió d'emocions amb emoticones i altres elements es poden interpretar com a signes clars de la proximitat i l'oralitat i es poden trobar en les dades com a característiques de CMC (Crystal, 2011: 48). D'altra banda, la prosòdia en Twitter només es pot tractar de forma limitada, per exemple pot ser indicada amb l'onomatopeïa o l'ús de majúscules (*op.cit.*: 20), tal com demostren els exemples (3), (4) i (5).

- (3) Com una floor que ebria de besos esclata PUM XIM PUM XIM...  
(Parlant PX, tuit 16148)

- (4) Una dona s'acaba de clavar dos dits sencers al nas amb la intenció de llevar – se un moc i NO ÉS BROMA. Espere que tinga sort i el puga treure (Parlant JJ, tuit 9869)
- (5) Necessite anar-me'n a la platja, JA! (Parlant MM, tuit 12849)

En definitiva, la comunicació en Twitter pot ser definida com un híbrid entre la llengua parlada i la llengua escrita. Es tracta d'enunciats escrits que semblen tenir més característiques de la llengua oral, o de la comunicació de la proximitat, que no pas característiques de la comunicació de la distància o de la llengua escrita (Crystal, 2011: 21). Amb tot, una classificació exacta només és possible si s'analitzen tuits publicats en la xarxa. En el cas del valencià, s'investiga en profunditat aquest aspecte en l'anàlisi de l'apartat 3.2. La finalitat és confirmar la hipòtesi que les publicacions de Twitter simulen la comunicació oral en forma escrita i poden ser interpretades, per tant, com a híbrids entre la comunicació oral i escrita, encara que això no es pot generalitzar a tota la comunicació generada en Twitter.

## ■ 2 Metodologia

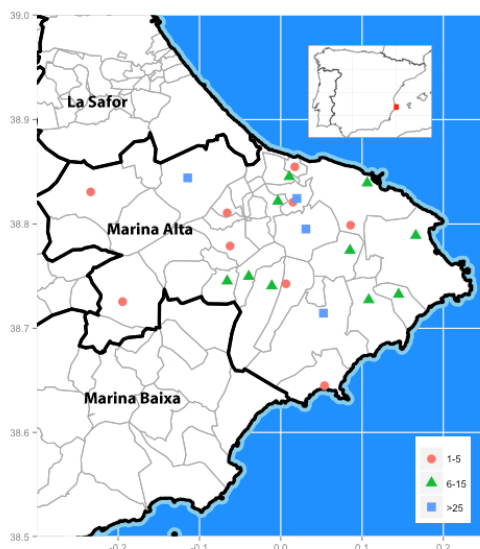
Després d'exposar el marc teòric d'aquesta recerca, es presenta la metodologia seguida per aconseguir l'objectiu del treball. L'estudi s'ha iniciat amb la creació d'un corpus format per missatges de la xarxa Twitter, fet que ha permès estudiar l'ús de la llengua.<sup>6</sup> En un període de cinc mesos, s'ha creat una base de dades amb 324.879 mostres, que són un total de 32.155 tuits de 241 usuaris. Els retuits, és a dir, els missatges reenviats, no s'han inclòs en el corpus perquè molt sovint no provenien dels usuaris seleccionats.

Pel que fa a la creació del corpus, s'ha utilitzat la tècnica del *data mining* (Han / Kamber, 2011). Això significa que amb el paquet *Twitter* (Gentry, 2009), disponible en el programa R (R Core Team, 2014), s'ha generat una API (*application program interface*) amb la qual s'ha establert una connexió entre el programa R i la xarxa Twitter. Així doncs, s'ha configurat de manera semiautomàtica el corpus i en un període de temps de cinc mesos s'ha pogut confeccionar una base de dades basada en els tuits dels usuaris seleccionats.

---

6 El corpus s'ha realitzat en el marc del projecte doctoral de l'autora, que es desenvolupa a la Universitat de Freiburg dins del grup de recerca DFG 1624 "Frequency effects in language" i es titula "Frequenzeffekte bei der Normierung romanischer Minderheitensprachen: Das Beispiel des Valencianischen" (en preparació).

Dins el marc d'aquest projecte, la xarxa Twitter s'ha definit com un text (Milner, 2011), és a dir, com un arxiu de llengua virtual. Les dades extretes són *screen-based*, això vol dir que són produïdes pels usuaris i recollides en línia pel científic (Androutsopoulos, 2013: 240), mentre que les dades que són *user-based* impliquen la interacció entre el científic i els usuaris (*ibid.*). D'acord amb els criteris de Herring / Barab / Kling (2004: 351–354), les dades de CMC s'han reunit seguint el criteri del mostratge individual o del mostratge en grup (*sampling by individual or group*). En conseqüència, s'han respectat criteris sociodemogràfics com per exemple la procedència dels usuaris i les seves edats. De vegades ha resultat complicat obtenir aquestes dades personals, per aquesta raó només s'han seleccionat els perfils d'usuaris actius en els quals l'edat dels parlants (entre 15 i 35 anys) i la procedència<sup>7</sup> eren evidents. Després, els seguidors dels subjectes escollits que tenien més o menys la mateixa edat i que mostraven al seu perfil la seva procedència també han estat inclosos en la base de dades. Així, ha estat possible establir relacions comunicatives entres els usuaris seleccionats. Per tant, s'ha pogut mostrar la comunicació d'una part de la comunitat lingüística valenciana en una xarxa social.



Mapa 1: Origen dels parlants en el corpus.

7 La informació sobre la procedència dels usuaris s'ha pogut deduir de la indicació del lloc geogràfic o de les coordenades exactes (*geotagging*) que apareixen en els perfils.

Pel que fa a l'origen dels parlants, s'ha limitat a usuaris que provenen de la comarca valenciana de la Marina Alta, província d'Alacant, com es pot veure en el mapa 1. El corpus està format per parlants dels següents pobles d'aquesta comarca: Alcalalí, Beniarbeig, Benissa, Benitatxell, Calp, Castell de Castells, Dénia, el Verger, Gata de Gorgos, Jesús Pobre, la Vall de Gallinera, Llíber, Ondara, Orba, Parcent, Pedreguer, Pego, Sagra, Sanet i Negrals, Teulada, Xàbia i Xaló.

Un altre aspecte que juga un paper molt important en aquesta investigació és la protecció de dades. Quant a això es pot dir que només s'han utilitzat perfils als quals es pot accedir lliurement, és a dir, que són perfils públics. A més, en aquest treball quan s'ha treballat amb citacions de tuïts s'han utilitzat sempre pseudònims per garantir la privacitat dels parlants i per assegurar el màxim anonimat.

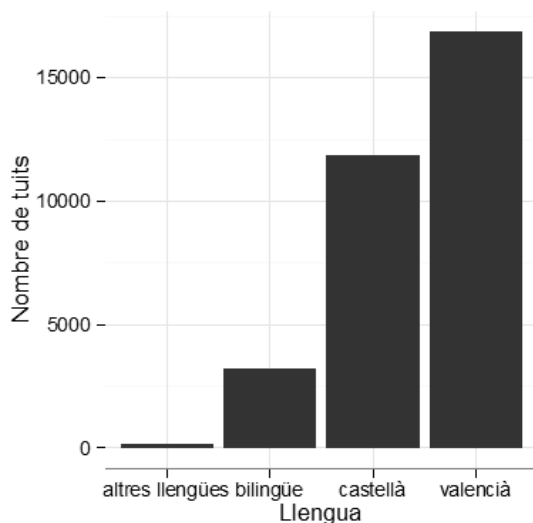
En resum, es pot dir que el corpus *Twitter* serveix com a reproducció de la llengua dels joves valencians procedents de la Marina Alta dins d'una xarxa social. Si es té en compte l'aspecte del valencià com a llengua minoritària, l'anàlisi del corpus *Twitter* esdevé molt interessant perquè els mitjans digitals solen ser classificats com a mecanismes importants per a la revitalització de les llengües minoritàries (cf. Jones, 2013: 58). A través de les xarxes socials com Twitter, però també de Facebook o altres xarxes, els parlants de la llengua minoritària es troben diàriament en contacte amb la seva llengua. Gràcies a la mediatització, la llengua pot experimentar una ampliació (Moring / Husband, 2007: 78) i, per tant, el seu ús es pot normalitzar. En el següent capítol s'aprofundeix en l'ús de la llengua en Twitter.

### ■ 3.1 Anàlisi I: L'ús del valencià i el castellà en Twitter

En aquest apartat es tracta el primer objectiu de la investigació: analitzar l'ús de les llengües en la comunicació de Twitter. Concretament, hem examinat quina de les llengües oficials de la Comunitat Valenciana és la més utilitzada pels tuïtaires per poder així saber si el valencià és una llengua viva en la xarxa social Twitter.

En examinar quantitativament els tuïts del corpus *Twitter* s'ha observat que d'un total de 32.155 tuïts, uns 16.881, és a dir, més de la meitat dels tuïts, exactament un 52,5%, estan escrits completament en valencià, mentre que un total d'11.897 tuïts, un 37%, estan escrits en castellà. També cal destacar que s'han trobat elements bilingües (valencià–castellà) en 3.215 tuïts (10%). A més, unes 160 entrades (0,5%) estan escrites en altres llen-

gües (sobretot en anglès), o bé la llengua del tuit no és identificable. Tal com demostra el gràfic 1, es veu una clara tendència a utilitzar el valencià en les comunicacions produïdes en Twitter.



Gràfic 1:  
Nombre de tuits per  
llengua en el corpus  
*Twitter*.

A banda del fet comprovat que la majoria dels tuits estan escrits en valencià, el bilingüisme és un aspecte especialment interessant que caracteritza la realitat sociolingüística de la Comunitat Valenciana, tot i que els tuits bilingües només representen un 10% del corpus. En els tuits identificats com a bilingües, els joves utilitzen les dues llengües (valencià i castellà) i canvien la llengua dins d'un mateix tuit. En aquest cas es poden produir dos tipus de canvis. D'un costat el canvi pot ser intrafrasal, com veiem en l'exemple (6):

- (6) @AidaColom bon diaaa!!! Auras descansat eee?? Jajajaa yo tinc una soon!!! Me vas abandonar en alma, sols stava tu cuerpo serrano!!  
(Parlant KD, tuit 29)

L'usuari en l'exemple (6), a banda d'utilitzar la grafia castellana, com per exemple en *yo*, que també podria ser interpretat com a element castellà, canvia completament al castellà al final de la frase (*tu cuerpo serrano*), fet que pot ser considerat com a exemple d'un canvi intrafrasal.



D'un altre costat, el canvi de llengua en un tuit bilingüe pot ser interfrasal, tal com ho demostra l'exemple (7):

- (7) @ArantxaPrz vale ya me paso q creo q el miércoles ya stoy d vacas jijiji encara tens internet?? (Parlant TF, tuit 93)

Les dues frases en el tuit de l'exemple (7) no estan exactament segmentades per un punt, però l'onomatopeia *jijiji* demostra que s'acaba la primera frase i, per tant, es pot definir aquest canvi com a interfrasal, ja que després el parlant continua el seu tuit en valencià (*encara tens internet?*).

El fet que els joves generin tuits bilingües reflecteix la seva realitat lingüística diària, però el nombre limitat de tuits bilingües demostra també que en certa manera els parlants tendeixen a distingir entre les dues llengües, és a dir, donen funcions diferents a cada llengua a l'hora de comunicar-se en la xarxa. Per tant, caldria aprofundir en quines estructures comunicatives apareix el fenomen de tuits bilingües. Això també significa que, per exemple, s'haurien d'examinar més detalladament les estructures lingüístiques dels seguidors dels parlants en el corpus *Twitter* per veure en quines situacions, o sigui comprovar amb quins usuaris, els parlants canvien al castellà. Tanmateix, aquest objectiu excedeix l'abast d'aquesta recerca i es portarà a terme en un proper estudi.

En resum, es pot afirmar que en el corpus analitzat existeix un gran nombre de tuits escrits completament en valencià, encara que també es poden trobar, en una quantitat més baixa, tuits en castellà, tuits bilingües i fins i tot tuits en els quals la llengua no s'ha pogut identificar. No obstant això, el fet que la majoria de tuits siguin en valencià demostra que aquesta llengua gaudeix d'un ús normal entre els parlants. En definitiva, el resultat obtingut en aquesta anàlisi demostra la vitalitat del valencià en una comarca com la Marina Alta on la presència de llengües estrangeres, com a conseqüència de la immigració europea i el turisme, fan més difícil l'ús del valencià en determinats àmbits.

### ■ 3.2 Anàlisi II: L'ús de diferents varietats de la llengua

Després de comprovar que els parlants del corpus utilitzen generalment més el valencià que no pas el castellà, el següent pas ha estat examinar quina varietat de la llengua fan servir els usuaris quan tuitegen.

Per refutar la hipòtesi anteriorment comentada sobre el caràcter oral de la comunicació en la xarxa social Twitter, que sembla aparentment un mitjà

de comunicació basat en la llengua escrita, en aquest apartat s'exposa si els tuits examinats en l'estudi contenen un major percentatge d'elements de la llengua oral o, per contra, de la llengua escrita. Concretament, s'ha determinat si els parlants utilitzen característiques típiques i prominents del valencià que poden ser considerades sovint com a elements clarament col·loquials, perquè no estan integrades en la normativa oficial valenciana, o si, en canvi, fan servir elements de la llengua escrita que corresponen normalment amb formes del català estàndard, basat en el català oriental, o amb formes que també són acceptades per la normativa oficial valenciana.<sup>8</sup>

En relació amb l'aspecte de la normativa, cal mencionar que existeixen diferents normes prescriptives en el territori de parla valenciana, de les quals només una, la de l'Acadèmia Valenciana de la Llengua (AVL), és considerada com la norma oficial valenciana (cf. Estatut d'Autonomia, 2006, Art. 6.8). En les normes existents es fan patents les ideologies de les institucions normatives. Es podria dir que hi ha tres ideologies representades en tres normes diferents en el territori de parla valenciana. A continuació, es presenten breument les tres institucions normatives.

Com ja s'ha mencionat, la institució en matèria lingüística de la Comunitat Valenciana és l'Acadèmia Valenciana de la Llengua. Fundada l'any 1998, la seva tasca és tant elaborar com promoure la normalització lingüística del valencià, que inclou també la seva normativització basant-se en les normes de Castelló (GVA, 1998). L'AVL defensa la unitat de la llengua catalana i defineix el valencià com a:

llengua romànica parlada a la Comunitat Valenciana, així com a Catalunya, les Illes Balears, el departament francès dels Pirineus Orientals, el Principat d'Andorra, la franja oriental d'Aragó i la ciutat sarda de l'Alguer, llocs on rep el nom de català. (AVL, 2014)

La institució ha publicat diferents documents normatius com la *Gramàtica Normativa* (AVL, 2006a), el *Diccionari Ortogràfic i de Pronunciació del Valencià* (AVL, 2006b) i, darrerament, de moment només en una versió en línia, el *Diccionari normatiu valencià* (AVL, 2014), que respecta les normes de Castelló de l'any 1932 (AVL, 2005) basades en les normes de Pompeu Fabra de principis del segle XX, però adaptades a les particularitats de la varietat valenciana (Ferrando Francés / Nicolás Amorós, 2011: 365).

---

8 Pel que fa la classificació dialectal en els territoris de parla catalana, veieu per exemple Veny (2007).

Una altra institució normativa és l'Institut d'Estudis Catalans (IEC). La seva normativa pren com a base les normes de Pompeu Fabra i defensa la unitat de la llengua catalana (cf. IEC, 2013). L'abast de la normativa emesa per l'IEC compren tots els territoris on es parla català, és a dir, respecta totes les varietats del català i intenta tenir en compte les variants dialectals en els seus documents normatius com, per exemple, en el *Diccionari de la Llengua Catalana* (IEC, 2007) i la *Gramàtica Normativa* (IEC, 2014). La normativa de l'IEC no ha estat considerada la normativa lingüística oficial per part del govern de la Generalitat Valenciana, però aquesta és semblant a la normativa de l'AVL perquè les dues es basen, en certa manera, en les normes de Pompeu Fabra, que en el cas de la normativa oficial valenciana s'han adaptat en les Normes de Castelló (GVA, 1998). A més, les dues normatives tenen un caràcter convergent (cf. Gifreu, 2014: 175ss.).

Finalment, cal esmentar una altra institució normativa que ha pretès normativitzar el valencià, la Real Acadèmia de Cultura Valenciana (RACV), la qual segueix les normes del Puig creades l'any 1979 per codificar el valencià (cf. Àrias i Burdeos, 2006: 15). La RACV nega la unitat de la llengua i refusa adaptar la normativa catalana basada en les normes de Pompeu Fabra (cf. Andrés Durà, 2014). Fet que demostra la negació, per part de la RACV, del català i del valencià com a varietats de la mateixa llengua.

Tenint en compte aquestes tres institucions normatives (AVL, IEC i RACV), es pot parlar de tres models normatius (cf. Mas Castells, 2008: 123ss.). L'IEC representa un model uniformista i convergent que reconeix la unitat de la llengua catalana. La institució normativa oficial de la Comunitat Valenciana, l'AVL, constitueix un model convergent i particularista que també defensa la unitat de la llengua, però té en compte les particularitats del valencià. En darrer lloc, la RACV representa un model secessionista basat en la independència lingüística del valencià respecte al català (*ibid.*).

L'AVL és la institució oficial de referència de la llengua al territori valencià. Per tant, cal comprovar si els joves usen aquesta normativa en les xarxes socials com Twitter.

Per desenvolupar la investigació, s'han seleccionat variables lèxiques i morfosintàctiques amb variació dins dels diferents models normatius. Un primer exemple és el verb *eixir* que està considerat com un tret característic i prominent del valencià.<sup>9</sup> Aquest verb es troba en els documents norma-

---

9 Les dades sobre el grau de prominència de les variants mencionades provenen del projecte doctoral de l'autora (Tölke, en preparació).

tius de les tres institucions (IEC, AVL i RACV). L'IEC, però, marca el verb *eixir* com una variant del verb *sortir*, el qual no es considera com una forma típicament valenciana, ja que es fa servir, sobretot, per gran part dels catalanoparlants excepte els del territori valencià i de l'Alguer (cf. Veny, 2007). Aquesta variant (*sortir*) no està recollida en els documents normatius de la RACV (cf. López Verdejo, 2010); d'aquesta manera es pretén subratllar la diferència entre el català i el valencià i marcar el caràcter autònom del valencià. En el quadre 1 es mostren altres variables lèxiques i morfosintàctiques amb variació que formen part de les característiques examinades en aquest estudi.

no.	variable	variant	IEC	AVL	RACV
1	<i>eixir</i> <i>sortir</i>	VAL CAT	✓ variant ✓	✓ ✓ variant	✓ ✗
2	<i>hui</i> <i>avui</i>	VAL CAT	✓ variant ✓	✓ ✓ variant	✓ ✗
3	<i>Preposició en davant topònims</i> <i>Preposició a davant topònims</i>	VAL CAT	✗ ✓	✗ ✓	✓ ✗
4	<i>aixina</i> <i>així</i>	VAL CAT	✓ variant ✓	✓ variant ✓	✓ ✓ variant

Quadre 1: Variables amb variació i la seva acceptació en les normatives.

Per a cada variable del quadre 1, s'ha analitzat quines variants figuren com a acceptades en els documents normatius de les tres institucions lingüístiques estudiades. Tot seguit, s'exposen els resultats d'aquestes variables. La variable 1 (*eixir* / *sortir*) ja ha estat explicada anteriorment. La variable 2 conté els adverbis *hui* i *avui*. L'adverbi *hui* es considera la variant típica valenciana. En el quadre es veu que la forma *hui* està acceptada en dues institucions normatives valencianes, l'oficial de l'AVL i l'extraoficial de la RACV. En la normativa de l'IEC, la forma es considera com a variant de la forma *avui*. La variable 3 fa referència al cas de la preposició davant de topònims. La variant prominent valenciana és la preposició *en* davant d'un nom propi de lloc, p.ex. *Visc en Alacant*, encara que aquesta variant només és acceptada per la RACV. La preposició *a* davant topònims, és a dir, *Visc a Alacant*, és la variant acceptada per les institucions de l'AVL i l'IEC. En darrer lloc, en la variable 4 s'estudien les formes *aixina*, com a variant prominent valenciana, i *així*, com a variant catalana. La forma valenciana està

integrada en la normativa secessionista de la RACV i apareix com a variant en les normes de l'AVL i l'IEC.

És una conclusió lògica que totes les variants prominents valencianes com el verb *eixir*, l'adverbi *hui*, la preposició *en* davant topònims i l'adverbi *aixina* estiguin integrades en la normativa de la RACV, és a dir, la normativa secessionista, perquè la seva intenció és remarcar les diferències entre el valencià i el català. Les variants catalanes com el verb *sortir*, l'adverbi *avui*, la preposició *a* davant de topònims i l'adverbi *així* es troben totes en la normativa de l'IEC, que subratlla la unitat de la llengua catalana. A més, tots els trets valencians estan acceptats com a variants al costat de les formes catalanes. Finalment, l'AVL, la normativa oficial valenciana, accepta les variants *eixir* i *hui* i marca les correspondències amb *sortir* i *avui* com a variants. Això demostra l'actitud particularista d'aquesta normativa que d'un costat confirma la unitat del català però de l'altre té també la intenció de subratllar les particularitats de la varietat valenciana. L'AVL, però, no accepta l'ús de la preposició *en* davant de topònims. A més, marca la forma *aixina* com a variant col·loquial de l'adverbi *així*.

Un major aprofundiment en el corpus *Twitter* ha permès observar quin és l'ús de les quatre variables amb variació descrites anteriorment. Aquesta informació es presenta en el quadre 2, en què es mostren les freqüències absolutes i relatives d'una mostra de variables amb variació que s'observen en el corpus *Twitter*.

no.	variable	freqüència absoluta	freqüència relativa
1	<i>eixir</i>	140	0,95
	<i>sortir</i>	8	0,05
2	<i>hui</i>	627	0,87
	<i>avui</i>	93	0,13
3	Preposició <i>en</i> davant topònims	115	0,37
	Preposició <i>a</i> davant topònims	197	0,63
4	<i>aixina</i>	92	0,51
	<i>així</i>	88	0,49

Quadre 2: Freqüències en el corpus *Twitter*.

En general, si s'examinen les freqüències de les variants (cf. quadre 2 i gràfic 2), es pot veure que en el corpus *Twitter* s'utilitzen tant les formes considerades formes particulars valencianes, com les formes catalanes. Si

s'observen més en detall les variables 1 i 2, es comprova que les formes prominents valencianes (*eixir* i *hui*) són utilitzades més sovint que les corresponents catalanes (*sortir* i *avui*), tant si ens fixem en les freqüències absolutes com en les freqüències relatives. Per una banda, cal tenir en compte que els mots valencians estan admesos en les tres normatives, tot i que l'IEC accepta aquestes formes només com a variants. Per una altra banda, les variants catalanes (*sortir* i *avui*) són oficials en la normativa de l'IEC, però solament apareixen com a variants en la normativa oficial valenciana de l'AVL i no estan acceptades en la normativa secessionista de la RACV.

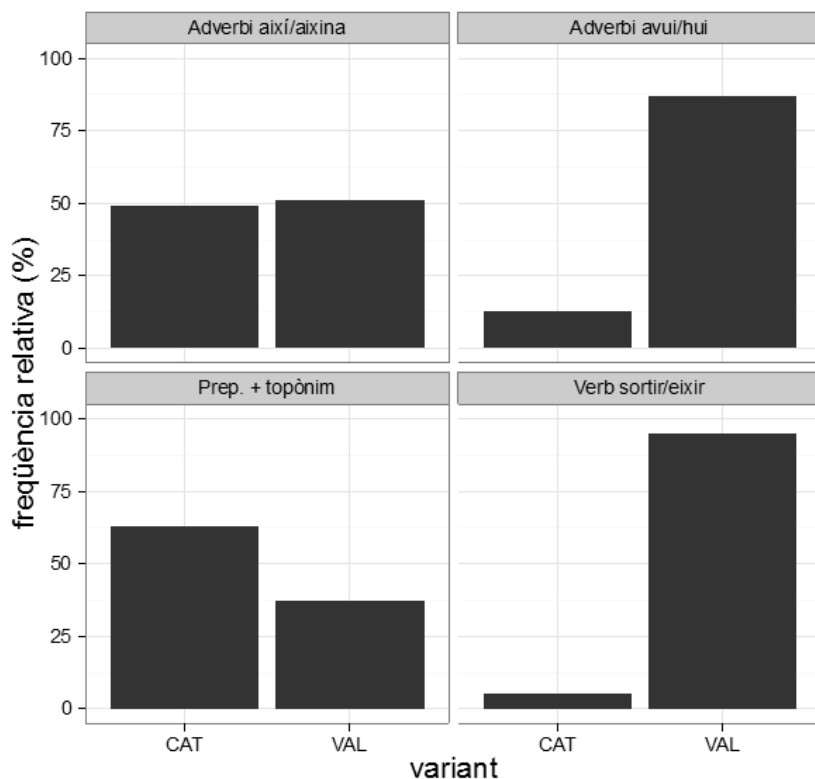
En examinar la variable 3, s'adverteix que la forma catalana (és a dir la preposició *a* davant de topònims) és la que s'utilitza més sovint (freqüència absoluta: 197, freqüència relativa: 0,63) pels parlants en el corpus estudiat. En canvi, la forma considerada com a típicament valenciana (preposició *en* davant de topònims) és usada amb una freqüència menor (freqüència absoluta: 115, freqüència relativa: 0,37). De fet, com podem veure en el quadre 1, només la normativa de la RACV accepta aquesta variant. Tant l'AVL com l'IEC no accepten la preposició *en* davant de topònims, sinó que és la preposició *a* l'única vàlida davant de topònim.

Pel que fa a la variable 4, es confirma que les variants *aixina* (freqüència relativa 0,51) i *així* (freqüència relativa 0,49) tenen una freqüència d'aparició molt semblant en el corpus *Twitter*. No obstant això, la variant valenciana *aixina* únicament està acceptada com a forma principal per la RACV, la qual considera l'adverbi *així* com a variant d'aquest mot. En canvi, l'AVL i l'IEC recullen el mot *aixina*, però com a forma col·loquial de l'adverbi *així*.

Si es comparen les freqüències d'ús de les variants (cf. gràfic 2) amb les acceptacions en les normatives, es pot dir que les formes que estan acceptades per la normativa oficial valenciana, és a dir, per l'Acadèmia Valenciana de la Llengua, són utilitzades amb major freqüència pels parlants en el corpus *Twitter*. Contràriament, les formes que la normativa oficial no accepta, com per exemple el cas de la preposició *en* davant de topònims, són menys utilitzades en el corpus *Twitter*.<sup>10</sup> A més, s'observa que els usuaris utilitzen les formes catalogades per l'AVL com a variants col·loquials, com és el cas de l'adverbi *aixina*, amb una freqüència semblant a la de l'adverbi *així*, considerat acceptació principal per l'AVL i l'IEC.

---

10 S'ha de tenir en compte igualment que alguns autors constaten que l'ús de la variant preposició *a* davant circumstancials de lloc és generalment molt més freqüent en la comarca de la Marina Alta que no pas en altres comarques (cf. Beltran i Calvo, 2011: 250; Colomina, 1985: 47).



Gràfic 2: Freqüència relativa de les variables amb variació en el corpus *Twitter*.

D'un costat, aquests resultats es poden considerar com una mostra de l'èxit de la normativa oficial del valencià de l'AVL, que se segueix en els centres educatius, tant als centres d'educació primària i secundària com a les universitats públiques de la Comunitat Valenciana. D'un altre costat, d'aquestes observacions es pot deduir que resulta complicat definir la comunicació en la xarxa Twitter. En el corpus, es poden trobar formes de la llengua oral i col·loquial que no es troben acceptades en cap model normatiu, amb excepció del model secessionista, que no és normatiu però es considera responsable de la normativització del valencià i intenta sempre potenciar totes les formes del valencià diferents del català, encara que aquests casos són relativament reduïts en el corpus analitzat. En les varia-

bles objecte d'aquest estudi, són les variants normativitzades les que tenen major presència. Per tant, es podria dir que són elements estandarditzades que pertanyen més a la parla escrita que no pas a la parla oral.

En resum, atès el resultat obtingut es pot afirmar que la definició exposada en l'apartat 1, que considera la comunicació en la xarxa Twitter com un híbrid entre la llengua parlada i la llengua escrita, és totalment justificada, perquè en la comunicació tenim tant elements de la llengua oral com de la llengua escrita.

#### ■ 4 Conclusió

Per concloure l'anàlisi del corpus *Twitter*, es pot afirmar que el valencià és la llengua d'ús habitual per part dels tuitaires joves de la Marina Alta que tenen entre 15 i 35 anys. De fet, l'anàlisi del corpus ha permès comprovar que es produeixen més tuits en valencià que tuits en castellà. A més, els tuits bilingües, és a dir, en castellà i valencià, es produeixen de manera molt reduïda en els usuaris estudiats. Aquest comportament lingüístic pot ser considerat com una prova de l'ús normal del valencià en la xarxa Twitter. Les xarxes socials esdevenen un reflex de la vida real d'aquesta comarca on la població utilitza el valencià en tots els àmbits de la vida, igual com ocorre amb parlants de qualsevol altra llengua, situació que contribueix a la normalització del valencià. Un altre element que indica una normalitat en l'ús del valencià és que només 37 dels 32.155 tuits totals analitzats en el corpus són publicacions que tracten sobre la situació del valencià, és a dir, s'adverteix que sols en un nombre escàs de tuits els usuaris parlen sobre el valencià i el conflicte lingüístic, com és el cas de l'exemple (8). Aquest exemple de tuit tracta indirectament de la definició i la denominació complicada del valencià i d'altres varietats catalanes, com el català parlat a la comunitat autònoma d'Aragó, i el fet que tots els catalanoparlants es poden entendre sense problemes perquè el català està constituït per un conjunt de varietats que tenen una mateixa base.

- (8) M'està sorprenent que al Congrés dels Diputats estan parlant constantment en Valencià!! O és català? Iber? LAPAO? Serà vlc xq ho entenc tot! (Parlant JJ, tuit 24667)

La resta dels tuits estudiats contenen temes diversos que no tenen res a veure amb la llengua, sinó que es relacionen amb temes quotidians. Aquest fet també és una mostra de la situació normalitzada de la llengua. En el cas



dels subjectes estudiats, el valencià s'utilitza de manera habitual i sense por perquè els joves es relacionen amb el seu entorn diàriament amb la seva llengua, fet que contribueix a la utilització normal del valencià en les xarxes socials, com Twitter.

Comptat i debatut, en aquest estudi s'ha demostrat que el valencià està ben viu a la comarca de la Marina Alta i que ha traspassat les fronteres d'intercanvi físic d'informació per a utilitzar-se en un espai virtual d'abast internacional en què s'equipara a l'estatus d'altres llengües. D'acord amb les paraules del guionista i comunicador valencià Eugeni Alemany en una entrevista en el diari digital *La Veu del País Valencià* (2014):

El futur del valencià està en les xarxes socials (...) la normalització del valencià passa per les xarxes (...) perquè amb les xarxes socials, el que dius s'escriu, és llegit per altres persones i això significa un aparador molt bestial per veure la presència de la nostra llengua. (Part, 2014) ■

## ■ Referències

- Acadèmia Valenciana de la Llengua (2005): *Dictamen sobre els principis i criteris per a la defensa de la denominació i l'entitat del valencià*, València: Acadèmia Valenciana de la Llengua, <<http://www.avl.gva.es/va/acordsAVL/main/03/document/NOMENTITAT.pdf>> [25.10.2015].
- (2006a): *Gramàtica normativa valenciana*, València: Acadèmia Valenciana de la Llengua.
- (2006b): *Diccionari ortogràfic i de pronunciació del valencià*, València: Acadèmia Valenciana de la Llengua.
- (2014): *Diccionari normatiu valencià*, València: Acadèmia Valenciana de la Llengua, <<http://www.avl.gva.es/dnv>> [25.10.2015].
- Alemany, Eugeni (2014): “Comunitat Twittera Valencianoparlant”, dins: *Ultramarinos Alemany*. Blog personal <<https://ultramarinosalemany.wordpress.com>> [25.10.2015].
- Andrés Durà, Raquel (2014): “¿Quién divide el catalán? El caso valenciano”, *La Vanguardia*, 25.09.2014, <<http://www.lavanguardia.com/local/valencia/20140925/54415385961/quien-divide-catalan-caso-valenciano.html>> [25.10.2015].

- Androutsopoulos, Jannis (2013): “Online data collection”, dins: Mallinson, Christine / Childs, Becky / Van Herk, Gerard (eds.): *Data collection in sociolinguistics: Methods and applications*, New York: Routledge, 236–251.
- Àrias i Burdeos, Juli Amadeu (2006): *Els verps en Llengua Valenciana i la seua flexió*, València: Real Acadèmia de Cultura Valenciana.
- Araeconomia @ARAEconomia, <<https://twitter.com/ARAEconomia/status/558596596943056896>> [25.10.2015].
- Beltran i Calvo, Vicent (2011): *Atles lingüístic de la Marina Alta*, Ondara: MACMA.
- Colomina i Castanyer, Jordi (1985): “El parlar de la Marina Alta”, *La Rella*, 37–54.
- Corporació Catalana de Mitjans Audiovisuals (CCMA) (2014): “Dos anys de twitter en català, però només 160.000 tuitaires”, <<http://www.ccma.cat/324/Dos-anys-de-twitter-en-catala-pero-nomes-160000-tuitaires/noticia/2441339/>> [25.10.2015].
- Crystal, David (2011): *Internet linguistics: a student guide*, London: Routledge.
- Doppelbauer, Max (2006): *València im Sprachenstreit. Sprachlicher Sezessionsismus als sozialpsychologisches Phänomen*, Wien: Braumüller.
- Ferrando Francés, Antoni / Nicolás Amorós, Miquel (2011): *Història de la llengua catalana*, Barcelona: Editorial UOC.
- Generalitat Valenciana (GVA) (1998): “Llei 7/1998 de 16 de Setembre de Creació de l’Acadèmia Valenciana de la Llengua”, <[http://www.docv.gva.es/fileadmin/user\\_upload/docvpub/Compendios/ACADEMIA\\_VALENCIANA\\_DE\\_LA\\_LENGUA.pdf](http://www.docv.gva.es/fileadmin/user_upload/docvpub/Compendios/ACADEMIA_VALENCIANA_DE_LA_LENGUA.pdf)> [25.10.2015].
- (2006): “Llei Orgànica 1/2006, de 10 d’abril, de Reforma de la Llei Orgànica 5/1982, d’1 de juliol, d’Estatut d’Autonomia de la Comunitat Valenciana”, <[http://www.docv.gva.es/datos/2006/04/11/pdf/2006\\_4177.pdf](http://www.docv.gva.es/datos/2006/04/11/pdf/2006_4177.pdf)> [25.10.2015].
- Gentry, Jeff (2009): *twitteR: R based Twitter client*, <<http://CRAN.R-project.org/package=twitteR>> [25.10.2015].
- Gifreu, Josep (2014): *El català a l’espai de comunicació. El procés de normalització de la llengua als mèdia (1976–2013)*, Bellaterra: Servei de Publicacions de la Universitat Autònoma de Barcelona.
- Han, Jiawei / Kamber, Micheline (2011): *Data Mining: Concepts and techniques*, Oxford: Elsevier.

- Herring, Susan C. / Barab, Sasha / Kling, Rob (2004): “Computer-Mediated Discourse Analysis: An approach to researching online behavior”, dins: Barab, Sasha / Kling, Rob / Gray, James H. (eds.): *Designing for virtual communities in the service of learning*, New York / Cambridge: Cambridge University Press, 338–377.
- Institut d'Estudis Catalans (2007): *Diccionari de la llengua catalana*, versió en línia, Barcelona: IEC, <<http://dlc.iec.cat>> [25.10.2015].
- (2013): *Normes ortogràfiques. Presentació*, versió en línia, Barcelona: IEC, <<http://normesortografiques.espais.iec.cat>> [25.10.2015].
- (2014): *Gramàtica de la llengua catalana*, versió en línia, Barcelona: IEC, <<http://www.iecat.net/institucio/seccions/filologica/gramatica>> [25.10.2015].
- Java, Akshay / Song, Xiaodan / Finin, Tim / Tseng, Belle (2009): “Why we twiTTER: An analysis of a microblogging community”, dins: Zhang, Haizheng (ed.): *Advances in Web Mining and Web Usage. 9th International Workshop on Knowledge Discovery on the Web, WebKDD 2007, and 1st International Workshop on Social Networks Analysis, SNA-KDD 2007*, Berlin / Heidelberg: Springer, 118–138.
- Jones, Elin Haf Gruffydd (ed.) (2013): *Social media and minority languages: Convergence and the creative industries*, Bristol: Multilingual Matters.
- Kailuweit, Rolf (2009): “Konzeptionelle Mündlichkeit?! Überlegungen zur Chat-Kommunikation anhand französischer, italienischer und spanischer Materialien”, *PhiN: Philologie im Netz* 48, 1–19, <<http://web.fu-berlin.de/phin/phin48/p48t1.htm>> [25.10.2015].
- Kemp, Simon (2014): “Digital Statshot 003. A compendium of digital statistics”, Singapore: We are social, <<http://de.slideshare.net/wearesocialsg/we-are-socials-digital-statshot-003>> [25.10.2015].
- Koch, Peter / Oesterreicher, Wulf (2011): *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*, Berlin / New York: De Gruyter.
- López Verdejo, Voro (2010): *Diccionari general de la Llengua Valenciana*, València: Real Acadèmia de Cultura Valenciana, <<http://diccionari.llenguavalenciana.com/general/>> [25.10.2015].
- Mas Castells, Josep Àngel (2008): *El morfema ideològic. Una anàlisi crítica dels models de llengua valencians*, Benicarló: Onada.
- McQuail, Denis (2010): *McQuail's Mass Communication Theory*, London: Sage.

- Milner, Ryan (2011): “The study of cultures online: Some methodological and ethical tensions”, *Graduate Journal of Social Science* 8:3, 14–35.
- Moring, Tom/ Husband, Charles (2007): “The contribution of Swedish-language media in Finland to linguistic vitality”, *International Journal of the Sociology of Language* 187–188, 75–101.
- Part, Leo (2014): “Enquesta d’@EugeniAlemany sobre l’ús del #valencià a les xarxes socials”, *La Veu del País Valencià*, 2.11.2015, <<http://www.laveupv.com/noticia/11722/enquesta-deugenialemany-sobre-lus-del-valencia-a-les-xarxes-socials>> [25.10.2015].
- R Core Team (2014): *R: A language and environment for statistical computing*, Wien: R Foundation for Statistical Computing, <<http://www.R-project.org/>> [25.10.2015].
- TERMCAT (2015) <<http://www.termcat.cat/ca>> [25.10.2015].
- Tölke, Vanessa (en preparació): *Frequenzeffekte bei der Normierung romanischer Minderheitensprachen*, Freiburg im Breisgau: Albert-Ludwigs-Universität [tesi doctoral].
- Veny, Joan (2007): *Petit atlas lingüístic del domini català*, Barcelona: Institut d’Estudis Catalans.

- Vanessa Tölke, Albert-Ludwigs-Universität, Graduiertenkolleg DFG 1624 “Frequenzeffekte in der Sprache”, Belfortstraße 18, D-79098 Freiburg im Breisgau, <[vanessa.toelke@frequenz.uni-freiburg.de](mailto:vanessa.toelke@frequenz.uni-freiburg.de)>.

Zusammenfassung: Die Intention des vorliegenden Beitrags ist es, den Gebrauch des Valencianischen innerhalb des sozialen Netzwerks Twitter zu untersuchen. Es wird sich der Frage angenähert, ob Twitter zum Erhalt von Minderheitensprachen, wie dem Valencianischen, beitragen und somit als Stütze für die Normalisierung des Sprachgebrauchs gelten kann. In diesem Kontext wird ebenfalls versucht, die Kommunikation innerhalb von Twitter zu definieren und Aussagen darüber zu treffen, ob es sich bei der in Twitter benutzten Sprache um geschriebene oder gesprochene Sprache handelt. Dieser Fragestellung wird anhand einer quantitativ ausgerichteten Studie nachgegangen, indem ein auf dem Netzwerk Twitter basierendes Korpus analysiert wird. Zusätzlich wird untersucht, inwieweit sich die untersuchten Sprecher des Korpus beim Twittern an der offiziellen valencianischen Sprachnorm, vorgegeben durch die *Acadèmia Valenciana de la Llengua*, orientieren.

Die Analyse zeigt, dass das Valencianische die gebräuchlichste Sprache unter den Sprechern des Korpus ist, was als ein Nachweis für den normalisierten Gebrauch des Valencianischen zumindest in einer spezifischen Region der valencianischen Gemeinschaft, der Marina Alta, gelten kann. Außerdem kann die Hypothese bestätigt werden, dass sich die Kommunikation im Netzwerk Twitter aus einem Hybrid von gesprochener und geschriebener Sprache konstituiert. ■

Summary: This article's aim is to analyze the use of Valencian in the social network Twitter. It addresses the issue of whether Twitter can contribute to the preservation of minority languages such as Valencian and can thus be regarded as a support for the normalization of the language. In this context an attempt is made to define communication within Twitter and to establish whether the language used in the network is characterized as spoken or written language. These questions are explored via a quantitative approach, based on a corpus of tweets in Valencian. The article further examines to what extent Twitter users follow the official Valencian language norm, represented by the *Acadèmia Valenciana de la Llengua*.

The analysis reveals that Valencian is the most commonly used language amongst the speakers of the corpus, which could serve as evidence for the normalized use of Valencian at least in a specific region of the Valencian Community, the Marina Alta. Additionally, the hypothesis that communication in Twitter is composed of a hybrid of spoken and written language can be confirmed. [Keywords: Valencian Catalan; language conflict; norm conflict; computer-mediated communication] ■