



HAL
open science

Proximité/distance comme catégories d'analyse pour la communication médiée par ordinateur : l'exemple de Twitter en allemand

Laurent Gautier

► **To cite this version:**

Laurent Gautier. Proximité/distance comme catégories d'analyse pour la communication médiée par ordinateur : l'exemple de Twitter en allemand. Doctorat. Sorbonne Université - Centre Malesherbes, France. 2019. cel-02004562

HAL Id: cel-02004562

<https://shs.hal.science/cel-02004562>

Submitted on 1 Feb 2019

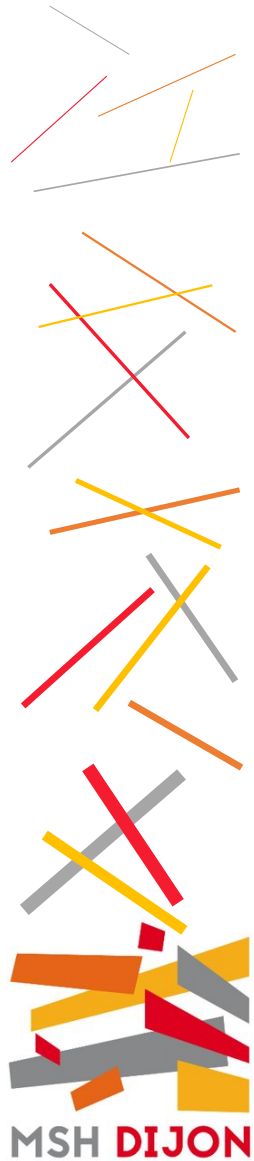
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Proximité/distance comme catégories d'analyse pour la communication médiée par ordinateur : l'exemple de Twitter en allemand

Laurent GAUTIER, Centre Interlangues Texte Image Langage (UBFC, EA 4182) &
MSH Dijon (USR CNRS – uB 3516)



Plan de la communication

1. Contexte
2. Problématique
3. Corpus
4. Proximité/Distance
 1. Une narration éclatée
 2. Linéarisation, cohésion et cohérence
5. Bilan et perspectives

Contexte

Nouveaux usages, nouveaux objets de recherche pour le linguiste

- Les données numériques **natives** comme nouveaux corpus (Schmitz 2004, Longhi 2012, Paveau 2013, 2015) => « écologie du discours numérique » :
 - Nouveaux types de discours analysés / d'acteurs / d'interactions
 - Facilité d'accès trompeuse (droit, technique), *cf.* section 3
- Les **réseaux sociaux** comme nouveaux objets de recherche transdisciplinaire :
 - Communication médiée par ordinateur (**CMO**) (Herring / Stein / Virtanen 2013)
 - Approche quali traditionnelle facilement doublée par quanti (Guilbert 2014, HS de Corela)

Contexte

Nouvelles formes d'écriture

(Liénard 2011, 2012 ; Dürscheid/Frick 2016)

- Poids des **dispositifs** discursivo-techniques sur les pratiques d'écriture (sms, Reinkemeyer 2013 ; whats'app, Dürscheid 2018a,b) => *Technolinguism* (Pfrehm 2018) :
 - Terminal
 - Émoticônes
 - Saisie intuitive
- Poids de ces mêmes dispositifs sur le résultat de l'écriture
- Conséquences pour l'acte de décodage et de **construction** du sens
- **Brièveté** comme caractéristique clef : 140 (*2) caractères ici

Contexte

Une double « digitalisation » que la recherche en linguistique ne peut ignorer

- La CMO comme objet linguistique *sui generis* :
„<...> durch Digitalisierung und Vernetzung wird Sprache anders dokumentierbar und erforschbar, sie wird in anderer Art zu sprachlichen Äußerungen verdichtet, und **mittels Computern wird auf andere Weise im Internet sprachlich kommuniziert.**“ (Lobin 2018 : 7, souligné LG)
- De nouvelles méthodes et de nouvelles catégories indispensables :
„Der Digitalisierung der Sprachverwendung entspricht eine **Digitalisierung ihrer Erforschung.**“ (Lobin 2018 : 9)

Problématique

Les corpus de tweets : pour des approches intégrées

- Approche intégrée nécessite une approche **(micro)linguistique** des tweets (vs majorité des approches SIC) :
 - considérés dans leur face signifiante ;
 - envisagés comme **'micro-textes'** (donc soumis aux / à des règles de textualité) ;
 - insérés dans un dispositif sociotechnique interagissant avec des formes de communication plus traditionnelles (Siever/Schlobinski 2013, Thimm sous presse, Androutsopoulos/Busch sous presse)

Problématique

Double problématique :

Théorique :

- Quels sont les **impacts** du dispositif sociotechnique sur le mode d'écriture du message (=> *technolinguism*) ?
- Quelles conséquences doit-on en tirer pour l'appréhension de la **textualité** et de l' **interactionnalité** du tweet ?

=> Questions testées ici à partir des oppositions écrit/oral-proche/distant

Pratique :

- Comment les scripteurs envisagent et gèrent-ils ces aspects dans un cadre spatialement **contraint** et fonctionnellement **prédéfini** (opérateurs) ?

⇒ **Quelle littératie numérique pour le locuteur-scripteur de la CMO ?**

Problématique

« (...) la **littératie numérique** n'est **pas une catégorie technique** qui décrit un niveau fonctionnel minimal de compétences technologiques, mais plutôt une vaste capacité de **participer à une société qui utilise la technologie des communications numériques** dans les milieux de travail, au gouvernement, en éducation, dans les domaines culturels, dans les espaces civiques, dans les foyers et dans les loisirs ».

(Hoechsmann / DeWaard 2015 : 5)

Problématique

Adopter une approche scientifique et non naïve

- Pour sortir de la *laienlinguistische Sprachkritik* : „Kritik an bestimmten Sprachnormen (...), die von Nicht-Linguisten und ohne empirisch gesicherte Untersuchung des Sprachgebrauchs geäußert wird.“ (Kilian/Niehr/Schiewe 2010: 56), exemple du # (Gautier/Michel, en prép.)
- Pour ne pas décontextualiser – **épistémologiquement et heuristiquement** – l’analyse : „Das Bild der Sprache prägt nicht nur die in einer Kulturgemeinschaft anerkannten Normen und Werte, es dokumentiert zugleich die öffentliche Diskussion darüber“. (Lobin 2018 : 7)

Gerade das Hashtag ist so praktisch, stellt es das Geschriebene doch gleich unter eine übergeordnete Kategorie, ohne dass man dazu weitere Worte der Erklärung (etwa „Soweit meine jüngsten Erlebnisse aus der Kategorie: Versagen im Alltag“) anbringen müsste. So praktisch, dass man dringend überlegen müsste, wie man das auch im Verbalsprachlichen zum Ausdruck bringen könnte.

rosenkrieger PERMALINK

9. Juli 2009 7:01 pm

Ich finde die Hashtags eine sehr gute „Erfindung“. Wenn man aber mal ehrlich ist sind es ja eigentlich „Tags“ 😊 Tags sind eine wunderbare Methode um Dinge zu sortieren (machen wir ja in Blogs mit den Tag Clouds nicht anders), bei Del.icio.us die Links etc etc.

Mit Hashtags/Tags in Mails gehts auch leichter. Ich hoffe das es sich durchsetzt und Mainstream wird.

Corpus

De nouvelles pratiques de collecte nécessaires

- Dimensions juridique et éthique :
« Twitter ne revendique aucun droit de propriété intellectuelle sur les contenus produits par les utilisateurs du service. (...) Mieux encore, Twitter encourage ses utilisateurs à verser les contenus par anticipation dans le domaine public ou à les placer sous licences libres pour en favoriser la réutilisation. » (Blog SI Lex de Lionel Maurel)
- Dimension **technologique** : compilation des données via l'API de twitter
- Dimension « **archivistique** » : gestion des métadonnées, structuration (xml/TEI)

Corpus

Un corpus original : tee2014

- MSH Dijon (TIL, Cimeos, LE2i) + Le Havre + Metz + partenaires dans 4 pays européens (Bonn pour GE) => 5 terrains nationaux, financements croisés (FNMSH, PHC Procope, CRB)
- Objet : communication « **générée** » par les candidats aux Elections Européennes de 2014 => 80 comptes par pays
 - Les messages envoyés sur les comptes Twitter des candidats
 - Les messages inclus dans les « conversations » entre ces comptes et d'autres tweetos (discours citoyens, débats internes...)
 - Les messages contenant les "hashtags" sélectionnés, liés à des thématiques politiques majeures de chaque pays
- 4 semaines de collecte : avant et juste après le scrutin

Corpus

Extraction du corpus global en allemand

- Corpus => lié aux comptes des candidats allemands (même si hétérogénéité linguistique) = 720.000 tweets (RT compris)
- Traitement pour interrogation (semi-)automatique
- Deux sorties : aspiration complète avec **méta-données** (sous SQL)+
texte brut

```
hopenina\\|Nina Hoppe\\|@_PaulSchmidt @MartinThuer dabei war das #tvduell sehr wohltuend. #umgang #sprache #respekt  
#haltung\\|  
JEF_de\\|JEF Deutschland\\|Toller Kommentar zum #tvduell im @zdf von @powl_b http://t.co/cIEfR69SYp #telleUROPE  
#Europatag\\|  
fun_pas\\|Pascal Funke\\|@tobiasraff Hehe, danke Deutschland :) #AfD\\|  
synee\\|Sven Clement\\|Komesch iwregens wei de Krop an d'AMCHAM meng Position zu #ttip komplett géigesätzlech  
interpretéieren ;)\\|  
FrauMaja\\|Maja Tiëgs, Pirat. \\|Leute, die bei #Piraten auch echt fehlen: @PjotrPopolskii Wo steckt der eigentlich?\\|  
HHerrnbrödt\\|Till\\|RT @Queeraten: Wenn ihr Vorschläge für Werbemittel bzgl. CSD habt, her damit: http://t.co/  
CJs5X198EN #Piraten^AM\\|  
seelenfarben\\|Annette\\|Es bleibt abzuwarten, was der #Rechtsruck bei der #Europawahl für den Tourismus bedeutet.  
#Tourismus Adieu Europa... #Europa\\|
```

```
from_userfrom_user_nametext  
@_PaulSchmidt @MartinThuer dabei war das #tvduell sehr wohltuend. #umgang #sprache  
#respekt #haltung  
Toller Kommentar zum #tvduell im @zdf von @powl_b http://t.co/cIEfR69SYp #telleUROPE  
#Europatag  
@tobiasraff Hehe, danke Deutschland :) #AfD  
"Komesch iwregens wei de Krop an d'AMCHAM meng Position zu #ttip komplett  
gÄ@igesÄ tzlech interpretÄ@ieren ;)"
```

- Interrogation sous AntConc

Proximité/distance

Privilégier le niveau conceptionnel

- Risque de confusion entre niveau du média et niveau de la conception, à l'origine de bien des approches naïves : cf. Söll (1974) repris et développé par Koch & Oesterreicher (2011 e.a.)

		Konzeption	
		<i>gesprochen</i>	<i>geschrieben</i>
Medium	<i>graphisch</i>	vergessen hab i	vergessen habe ich
	<i>phonisch</i>	[fɛə'gɛsən ha:bi]	[fɛə'gɛsən ha:bɔɪç]

Proximité/distance

- Interactions «orales » analysées à replacer le long d'un **continuum** « **proximité** communicationnelle » vs. « **distance** communicationnelle »

« Il y a donc, à côté de la dimension médiale (= le support : « écrit » [graph(ém)ique] vs.« oral » ou « parlé » [phon(ém)ique]), une dimension conceptionnelle : intuitivement, il semble qu'on ne tienne pas le même discours à l'écrit qu'à l'oral. Pourtant, évidemment, tout énoncé parlé peut être mis par écrit, et tout énoncé écrit peut être prononcé. **Les deux dimensions, médiale et conceptuelle, même si elles sont liées de fait dans l'usage, sont donc indépendantes de droit.** » (Modicom 2016 : 4)

Proximité/distance

Paramètres essentiels :

- Interaction privée/semi-privée/semi-publique/publique
- Degré de familiarité entre interlocuteurs
- Part laissée aux émotions (liée à l'objet de discours lui-même)
- Intégration de l'action et de la situation
- Co-Construction de la référence
- Proximité physique
- Coopération entre interlocuteurs
- Dialogicité
- Spontanéité
- Choix de l'objet de discours

Une narration éclatée (Thimm/Gautier 2014)



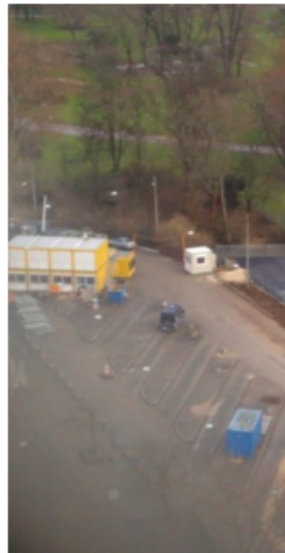
@stuttgarter1977
Andre Dietenberger

Zuerst so <http://twitpic.com/2tgfnx>
dann <http://twitpic.com/3wevim> dann
<http://twitpic.com/4025qg> zum Schluß
<http://twitpic.com/3wqbzp> #S21

vor 22 Minuten via web ☆ Als Favorit markieren ↻ Retweet ↩ Antworten
von Stammheim, Stuttgart



via TwitPic



via TwitPic



via TwitPic

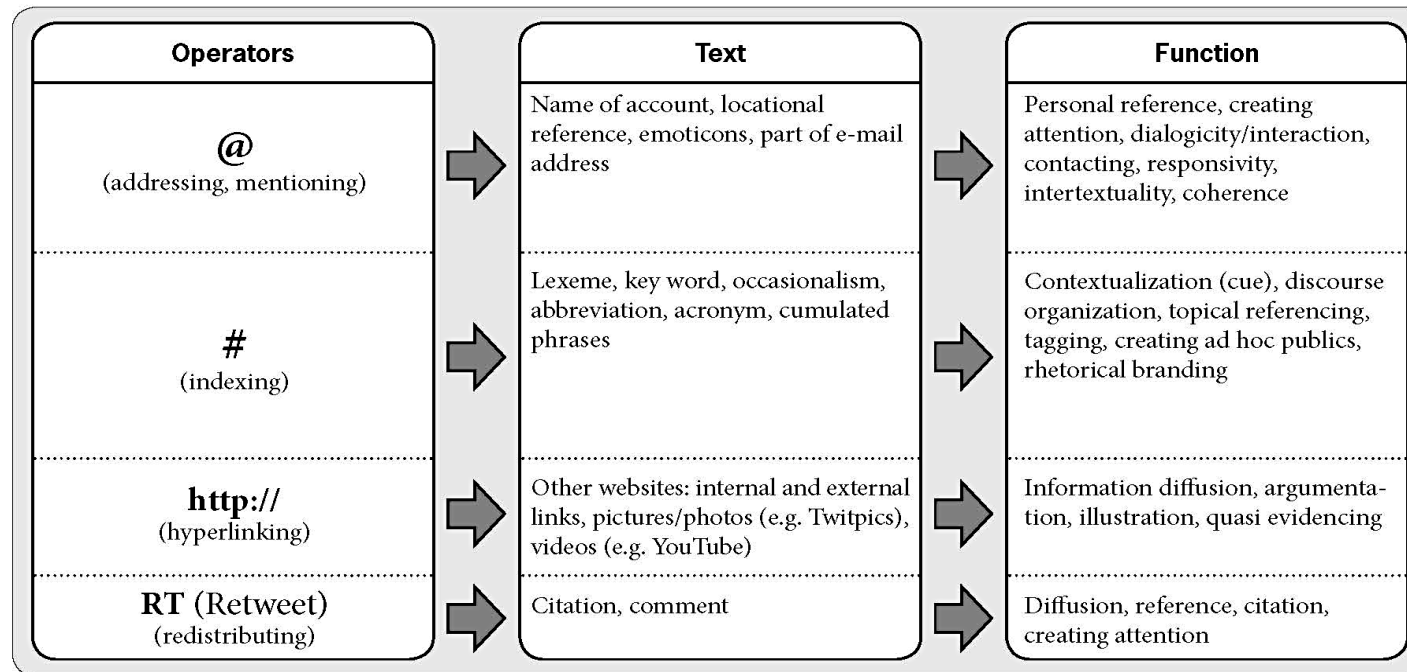


via TwitPic

Retweetet von [hering51](#) und 1 anderem

Une narration éclatée

Une conséquence directe des opérateurs discursivo-techniques



Thimm/Dang-Anh/
Einspänner (2012)

- Thèse défendue : ce sont des **facilitateurs de proximité** au niveau des contenus vs. **sources de distanciation** au niveau de la forme

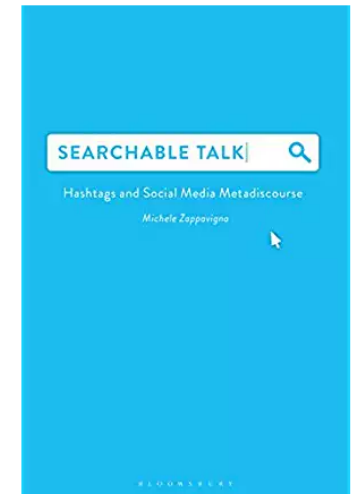
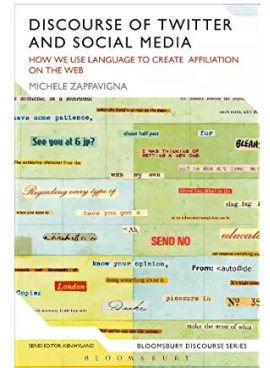
Linéarisation, cohésion et cohérence

- Remise en cause des principes de **linéarisation** comme base de construction de l'interaction :
 - (1) RT @PorcusDivinus: Mir etwas zu sehr auf "arme #afd-" gepolt, aber einige richtige, wichtige Ansätze dabei in der @SZ. <http://t.co/3K14pHF...>
=> Enjeu : garantir l'acceptabilité/lisibilité du tweet comme texte
- Réduction du degré d'informativité explicite à un minimum :
 - (2) Mir etwas zu sehr auf "arme #afd-" gepolt, aber einige richtige, wichtige Ansätze dabei in der @SZ.
=> Déplacement du lieu d'inscription de l'informativité
- Nécessité d'injecter dans le décodage outre la situationnalité de départ le savoir fonctionnel lié aux opérateurs @, http:// et #=>
Distribution complémentaire proximité/distance selon pôle communicationnel (production/réception)?

Linéarisation, cohésion et cohérence

et renégociation de la cohérence

- Double statut de #
 - # intégré syntaxiquement
- (3) @OlgaVanHorst @DMWarrior @BiboBissig Was konkret ist im #afd Wahlprogramm lächerlich? Ich finde es sehr fundiert!
- Décodé comme argument de l'énoncé (« niveau 0 ») = emploi mondain
- Décodé comme tête de paradigme = emploi autonymique



Linéarisation, cohésion et cohérence

- # non intégré syntaxiquement, en position d'ouverture (rare) et/ou de clôture (en masse)

(4) #AfD Stand vor dem Wittwer! Polizei hat bereits den gesamten #Schillerplatz umstellt! #stuttgart #bunt statt #braun #wk14 #reclaimeurope

- Fonction instructionnelle, « cadrative » impliquant/nécessitant un décodage segmenté
 - Participe à une cohérence de double niveau
 - Interne au texte
 - Supérieur au niveau du fil de tweets

=> Co-construction de la référence comme indice de proximité ?

Linéarisation, cohésion et cohérence

Cohérence interactionnelle

- Enjeu pour l'interprétation des opérateurs @ et RT
- **Enjeu pour la construction de la proximité, incidences sur:**
 - **Coopération**
 - **Dialogicité**
 - **Spontanéité**
- RT comme marqueur de diffusion, mais quid de la prise en charge énonciative surtout avec polyphonie exponentielle ?

(5) RT @FelixMSteiner: "Nicht #wählen die #AfD du darfst!" Danke @_verdi dafür!
#Europawahl #Humor <http://t.co/ogA6EkBang>

Linéarisation, cohésion et cohérence

- Double statut de @ comme marqueur d'adresse (9) et/ou mention (10) :

(6) @mattimerker @MartinSchulz: Danke Dir, klasse! Beste Grüße aus dem Wahlkampf in den Wahlkampf!

⇒ Décodage dépendant du remplissage textuel du tweet : vous d'adresse, acte de langage exclamatif + fonction phatique de merci

(7) Ich lach mich tot."#AfD - Auffangbecken für Dumme" Danke an @EinAugenschmaus für diese herrlich treffende Umbezeichnung.

⇒ Triple statut en contexte : mention + source + marqueur d'adresse

⇒ **Renforcement d'une proximité « marquée »**

Bilan et perspectives

- Tweet comme objet d'étude linguistique à part entière et extrêmement complexe :
 - différence **qualitative** avec d'autres formes d'écriture électronique (FB, blogs, forums)
 - enjeu **épistémologique** : construction d'une « théorie » du signe tweet revisitant des catégories pourtant anciennes : textualité, écrit/oral, recevabilité, cohérence et... proximité/distance
 - enjeu **méthodologique** : analyse quali devant inclure le balayage quanti de grandes masses

Bilan et perspectives

- La brièveté du tweet – contrainte spatiale – est la clef d'entrée pour l'analyse linguistique : textualité et potentiel interactionnel
- ⇒ Recours à une « **grammaire** » **des opérateurs** (Thimm sous presse)
- Ces caractéristiques le « prédisposent » à faire le **grand écart entre proximité et distance** selon les pôles production/écriture et réception/lecture-décodage/construction du sens
- ⇒ Limites de l'analyse linguistique stricte ⇒ glissement vers une **linguistique des usages** qui ne peut être qu'**empirique**



Merci pour votre attention !

Laurent GAUTIER, Centre Interlangues Texte Image Langage (UBFC, EA 4182)

& MSH Dijon (USR CNRS – uB 3516)

laurent.gautier@u-bourgogne.fr

