CERLIS

Centro di Ricerca sui Linguaggi Specialistici
Research Centre on Languages for Specific Purposes

Stefania M. Maci
& Michele Sala (eds.)

REPRESENTING AND REDEFINING SPECIALISED KNOWLEDGE: VARIETY IN LSP

8

Università degli Studi di Bergamo

CERLIS Series

Series Editor: Stefania M. Maci

Each volume of the series is subjected to a double peer-reviewing process.

CERLIS Series
Volume 8

Stefania M. Maci & Michele Sala (eds.)

Representing and Redefining Specialised Knowledge: Variety in LSP

2019
Università degli Studi di Bergamo

CERLIS SERIES Vol. 8

CERLIS
Centro di Ricerca sui Linguaggi Specialistici
Research Centre on Languages for Specific Purposes
University of Bergamo
www.unibg.it/cerlis

REPRESENTING AND REDEFINING SPECIALISED KNOWLEDGE:
VARIETY IN LSP
Editors: Stefania M. Maci, Michele Sala
ISBN 978-88-97253-03-7
ISSN 2532-2559 – CERLIS series [Online]
Url: http://hdl.handle.net/10446/155776
Doi: 10.6092/10446_978-88-97253-03-7

Contents

Adequacy, meaning construction and discursive conventions

STEFANIA CONSONNI

From Stigma to Statistics: A Study of US HIV Discourse in Digital Research Article Titles, 1986-2016[1]

1. Introduction

This chapter investigates Research Article Titles (RATs) published in medical journals worldwide from 1986 to 2016 and disseminating research on the Human Immunodeficiency Virus carried out in US health institutions and medical universities.

The clinical history of HIV in the US started in June 1981, when the deaths of five male homosexual patients in New York and Los Angeles from an unidentified syndrome, first labelled as 'gay cancer' or GRID (Gay-Related Immunodeficiency), were reported in the *Morbidity and Mortality Weekly Report* of the Centers for Disease Control and Prevention. The following two years, research was massively funded in both Europe and the US in order to identify the cause of a life-threatening and socially sensitive pathology, immediately associated with such controversial factors as high-risk sexual activity and injection-drug abuse. In September 1982, the name AIDS (Acquired Immune Deficiency Syndrome) was introduced, while illustrious casualties from the film, music and art industry attracted media speculation, thus casting an aura of stigma and scandal on social discourses about the disease. In May 1983, French virologists Françoise Barré-Sinoussi and Luc Montagnier at Institut Pasteur discovered LAV (Lymphadenopathy Associated Virus), the

[1] This study is part of a national research project on "Knowledge Dissemination across media in English: Continuity and change in discourse strategies, ideologies, and epistemologies", financed by the Italian Ministry of University and Research (PRIN 2015TJ8ZAS).

retrovirus responsible for AIDS; one year later, Robert Gallo's rival team at the National Cancer Institute announced the isolation of HTLV-III (Human T-Lymphotrophic Virus type III). In April 1985, the first World Health Organization conference on AIDS was held in Atlanta, Georgia; and in May 1986, HIV eventually entered the lexicon of pathology, as the International Committee on the Taxonomy of Viruses declared that the retrovirus triggering AIDS should be called HIV-1.

The present study investigates HIV research and discourse in the US – that is, the country with the world's highest funding for prevention and treatment programmes – along a time span of 30 years. Its purpose is to explore how and to what extent the language of expert-to-expert written medical communication has been changing with respect to two key epistemological factors:

1) The progressive advancement (and mutual influence), along the 1980s, 1990s and 2000s, of four clinical specialties, representing complementary standpoints as well as different steps forward in the global study, surveillance and containment of HIV. For reasons that will be specified, Public health, Molecular biology, Immunology and Infectious diseases are here taken as crucial research areas within the disciplinary evolution inherent to scientific knowledge about HIV/AIDS;
2) The methodological development, particularly from the early 1990s onwards, of a completely new set of procedures and protocols in the life sciences. As a matter of fact, the history of HIV appears to be intertwined with the inception of today's leading paradigm for medical knowledge, i.e., Evidence Based Medicine (EBM). EBM promotes systematic statistical analysis derived from the epidemiological scrutiny of large population samples as the best evidence available and the only possible criterion for the diagnosis and management of individual pathologies, with the purpose of minimizing bias and boosting expected result accuracy in a global perspective (Sackett *et al.* 1996; Greenhalgh 2010).

From a discourse-analytical and socio-semiotic angle, both of these factors point to the ideologically non-neutral nature of written scientific discourse, whereby research eminently depends on the

public dissemination of knowledge, a transactional phenomenon involving different actors and audiences, whose impact can hardly be distinguished from the scope and purposes of research itself (Bucchi 1998; Shinn/Whitley 1985). Knowledge validation is always negotiated among stratified audiences, including not only fellow- and training experts, but also parallel professional contexts, business and corporate audiences, and the lay public of popularization (Calsamiglia 2003; Calsamiglia/Van Dijk 2004; Garzone 2006; Gotti 2013; Myers 2003; Raffo 2016). Discursive feedback from all these strata contributes to science's legitimization, a process subtly impinging on the agendas of competing clinical methodologies and institutions, especially in the case of impactful pathologies such as HIV/AIDS. Even more so, the immense amplification brought about by the digital environment and by Web-based communication since the mid-1990s has been intensifying the influence of dissemination processes on scientific expository practices.

In this respect, the genre of medical RATs has evidenced itself as a crucial one. Influenced by media, advertising and entertainment discourse, and serving key audience-oriented purposes such as cognitive immediacy and emotional/aesthetic appeal (Hartley 2005, 2005b; Martin 1998; Straumann 1935), headlines have contributed to construing and legitimizing the institutional ideology of advanced industrial society, in both expert-to-expert and expert-to-layman contexts (Smith 2000; Soler 2007). Precisely depending on its conflation of pragmatic functions, the language of RATs has been deeply affecting the production and validation of scientific information (Garzone/Catenaccio 2008; Haggan 2004), from the conservative, gate-keeping formulations of traditional prose to more nuanced ones, placing emphasis on distinctively argumentative, persuasive and metadiscursive functions (Fortanet *et al.* 1998; Goodman / Thacker / Siegel 2001). The specialised contents of a medical RA (i.e., a text aimed at an expert audience and displaying the standard IMRD format, in order to convey the methods, hypotheses and results of experimental science) are in fact compressed and showcased in its title. A typically concise epitextual structure (Genette 1987), a RAT both synthetizes a corresponding article (in terms of informative content) and presents it in an efficient and appealing way,

providing the reader with directions in regard to the text's type and pragmatics.

In medical communication, RATs carry out a number of functions, the first and foremost of which is informativity (Smith 2000): they perform a "straightforward presentation of information" (Haggan 2004: 313), often being concise, complete and transparent advanced textual organizers, revealing preview information from a later and more extended text (Kozminksy 1977). Secondly, titles enable the indexing and search optimization of RAs, guiding the "organization and retrieval of scholarly data" (Soler 2007: 91) and surrogating them in reference lists, bibliographies, databases or the Web (Yitzhaki 1997). Finally, they are designed as to be attractive and persuasive with respect to the ensuing article: RATs do attract readers' attention, thus conditioning articles' selection on the part of the scientific community and/or other social actors involved in the dissemination process (Hjørland/Nielsen 2001). On the basis of the above, this study focusses on two research questions:

1) How has the language of medical RATs changed in the last thirty years in regard to, and across, the research specialties that have shaped the disciplinary evolution of HIV discourse, i.e., Public health, Molecular biology, Immunology and Infectious diseases? How, in turn, has the scientific scope and purpose of each specialty contributed to shape the language in which RATs disseminate medical knowledge?
2) To what extent, and in what ways, has the language of medical RATs responded to the implementation of a new overarching clinical methodology, i.e., Evidence Based Medicine (EBM)? What changes in the pragmatic scope and methodological positioning of written medical discourse about HIV has EBM brought about?

The next section will therefore be devoted to setting out the parameters by which this study will address such questions.

2. Materials and method

For the purpose of this analysis, a corpus of 4,504 RATs has been assembled, covering the totality of RAs presenting HIV research carried out in the US, published from 1986 to 2016 in 423 medical journals currently indexed within the Web of Science (WoS) Citation Index-Expanded.[2] The search was performed via the WoS Sci-Expanded analytic tools, setting the following criteria: HIV (Title field), English (Language), Research Article (Document type) and US (Country – of research, not of publication). The search yielded a total of 24,452 items, to which two further parameters for selection were applied:

1) The present corpus concentrates on seven sample (or apex) years, on the basis of 5-year intervals. Thus, only RATs published in 1986, 1991, 1996, 2001, 2006, 2011 and 2016 were collected. This resized the corpus to a total of 5,781 RATs.
2) A further analysis by Research area was carried out within each of the seven sample years. This evidenced four top-ranking specialties as shared across all sample years, and therefore as analyzable along the selected time span: Public health (479 RATs, representing 8.3% of the corpus), Molecular biology (444 RATs, 7.7%), Immunology (1,722 RATs, 29.8%) and Infectious diseases (1,859 RATs, 32.2%). Other specialties, covering the remaining 22.1%, were excluded from the corpus. A total of 4,504 RATs from 423 specialised journals were thus obtained and downloaded. Table 1 shows the distribution of RATs in the corpus (henceforth, WoS corpus), per year and per clinical specialty.

[2] Data were retrieved on 3 October 2018.

	1986	*1991*	*1996*	*2001*	*2006*	*2011*	*2016*	*Tot.*
1. Public health	0 (0%)	18 (32.1%)	13 (5.9%)	30 (7.4%)	46 (5.4%)	156 (7.8%)	216 (9.5%)	479 (8.3%)
2. Molecular biology	0 (0%)	15 (26.7%)	64 (29.2%)	58 (14.3%)	73 (8.5%)	128 (6.4%)	106 (4.7%)	444 (7.7%)
3. Immunology	0 (0%)	2 (3.5%)	33 (15%)	143 (35.4%)	343 (40.4%)	553 (27.8%)	648 (28.6%)	1,722 (29.8%)
4. Infectious diseases	0 (0%)	1 (1.7%)	19 (8.6%)	116 (28.7%)	355 (41.8%)	613 (30.8%)	755 (33.4%)	1,859 (32.2%)
TOT. RATs	0 (0%)	36 (64%)	129 (58.7%)	347 (85.8%)	817 (96.1%)	1,450 (72.7%)	1,725 (76.2%)	4,504 (77.9%)

Table 1. No. of RATs in the WoS corpus, per year and per clinical specialty.

From within a discourse-analytical framework, and assuming that titles carry out key pragmatic functions in regard to the informativity, retrievability and attractiveness of research articles (White/Hernandez 1991; Eyrolle / Virbel / Le Marié 2008; Hartley 2005, 2007), this study will quantify and analyse the evolution and variation patterns in the syntactic and textual construction strategies of 4,504 RATs. In order to do so, a preliminary charting of HIV's clinical coverage will be carried out in terms of the abovementioned four specialties (i.e., Public health, Molecular biology, Immunology and Infectious diseases), in order to define the disciplinary matrix (Bazerman 1988) within which HIV discourse is situated. On the basis of this, linguistic analysis will follow, focusing on the way medical RATs are worded out and organized at structural and textual level, with the purpose of investigating the impact of said specialties, on the one hand, and of Evidence Based Medicine, on the other, on the development of specialised communication about HIV. The analysis will be conducted on two levels:

1) At structural level (Fortanet *et al.* 1998; Haggan 2004; Yitzhaki 1997; Swales 2003; Soler 2007; Jaime Sisó 2009; White/Hernandez 1991; Hjørland/Nielsen 2001), RATs will be distinguished into conclusive, interrogative, nominal and compound; the distribution of titling constructions will be

studied within and across clinical specialties (research question 1).

2) At textual level, the use in RATs of expanded noun phrases bearing lexical reference to Evidence Based Medicine methodology (in particular in compound syntax) will be read as a metadiscursive strategy (Hyland 2005; Hartley 2007) textualizing EBM study design concerns, and evidencing changing attitudes towards the production and dissemination of medical knowledge in the 1980s, 1990s and 2000s (research question 2).

Materials have been scanned using the WoS Sci-Expanded analytic tools, AntConc (Anthony 2016) and WordSmith Tools (Scott 2017) software.

3. Results

3.1 Clinical coverage: No. of RATs per year and per specialty

Graph 1 (see Appendix) shows the distribution of RATs per year and per specialty, thus visualizing the main trends in the clinical coverage of HIV across the corpus.

3.1.1 Public health RATs

Public health RATs present research on the AIDS pandemics, focussing on risk *vs*. protective factors in the aetiology of the disease, with a view to protecting the health of local and global communities by means of "researching disease and injury prevention, and detecting, preventing and responding to infectious diseases" (CDC 2017). As shown in Graph 1 (see Appendix), Public health is the topic of 32.1% of RATs published in 1991 (18 occurrences), while its frequency sharply decreases as of 1996 (13 occurrences, 5.9%), never to peak again (7.4% in 2001, 5.4% in 2006, 7.8% and 9.5% in 2011 and 2016).

This seems to indicate that the initial coverage of HIV was mainly concerned with emergency issues such as epidemics-charting surveys and disease-control methods. This appears to be confirmed by such events as the introduction, in May 1987, of a travel ban issued by the US Public Health Services, whereby HIV was listed among the reasons for immigration exclusion, and testing became mandatory for all visa applicants.[3]

The ban, not to be lifted until January 2010, mirrored the early 'HIV exceptionalism' attitude of medical and political institutions facing the epidemic (Frieden *et al.* 2005), and went in concert with the social marginalization cast by sensationalism in the news media.[4]
The international effort to chart the infection through surveillance, estimates and monitoring (WHO 2018; UNAIDS 2018) – particularly in regard to high-risk patient categories – is evidenced in the WoS corpus by RATs such as the following:

(1) Behavioural, health and psychosocial factors and risk for HIV-infection among sexually active homosexual mean: The multicentre AIDS cohort study [1991]

(2) Preventing HIV/AIDS among high-risk urban women: The cost-effectiveness of a behavioural group intervention [1991]

---

3 The HIV/AIDS timelines used in this study were retrieved from HIV.gov (2016).

4 The early 1980s saw AIDS's first mention in national broadsheets and mainstream magazines: in July 1981, the *New York Times* published an article entitled "Rare Cancer Seen in 41 Homosexuals"; in August 1983, HIV became the leading story ("Gay America: Sex, Politics, and the Impact of AIDS") in the popular *Newsweek*. Only two years later did President Reagan publicly mention AIDS for the first time, tagging it as the country's top sanitary crisis. Only in the late 1980s and early 1990s did open signs of anti-stigmatization appear in the media, for instance when Diana, Princess of Wales, made headlines worldwide in April 1987 being photographed shaking hands with an HIV-positive patient in a London hospital; or when Tony Kushner's play about AIDS, *Angels in America*, was awarded the Pulitzer Prize for drama, in April 1993; and later that year, when the film *Philadelphia*, starring Tom Hanks, made an international case for HIV awareness.

(3) HIV-infection, genital ulcer disease, and crack cocaine use among patients attending a clinic for sexually-transmitted diseases [1991]

As shown by these examples, AIDS (i.e., the third and final stage of HIV infection) began to circulate in medical and media discourse as an aggressive pandemic, supposedly targeting only stigmatized social groups such as male homosexuals, male and female prostitutes and injection-drug users, while heterosexual, vertical (mother-child) and parenteral transmission (e.g. through blood transfusions, surgery etc.) remained substantially ignored until later in the decade.

3.1.2 Molecular biology RATs

Stemming from the functional study of cells at molecular level, this category of RATs presents experimental hypotheses mainly aimed at elucidating the biology (that is, the viral structure, and replication, transcription and translation processes) of HIV-1. As can be seen in Graph 1 (see Appendix), RATs in this specialty peak in 1991 and 1996, when they total, respectively, 26.7% and 29.2% (15 and 64 occurrences), while their frequency sharply decreases at the end of the 1990s, when they drop to 14.3% (2001), 8.5% (2006), 6.4% (2011) and 4.7% (2016). The descending trend may be due to the fact that these RATs tend to focus on two main top-priority topics in the 1990s, i.e., the pathogenesis and diagnosis of AIDS. On the one hand, the origination of the disease was at that time intensively studied: investigating the biochemical behaviour of its infectious agent was a prerogative for the understanding of the syndrome's biology.[5] On the other hand, early HIV diagnosis, the patenting of diagnostic tests and the screening of population samples was also a typical concern of

[5] As mentioned earlier, biochemical research in 1980s US mainly aimed, in competition with Europe, at identifying AIDS's cause and transmission. The 1983 dispute between Institut Pasteur and the National Cancer Institute as to who first discovered HIV did not end until 1987, when Ronald Reagan and Jacques Chirac agreed that the two countries should share credit for the discovery. In October 2008, however, the Nobel Prize for Medicine was awarded to Barré-Sinoussi and Montagnier, with only a passing mention of Robert Gallo's lab's achievements.

1980-90s Molecular biology research (CDC 1987).[6] Both issues are amply evidenced in the WoS corpus, as shown by the following examples:

(4) Post-PCR sterilization: Development and application to an HIV-1 diagnostic assay [1991]

(5) Identification of an ion channel activity of the Vpu transmembrane domain and its involvement in the regulation of virus release from HIV-1-infected cells [1996]

(6) Long-term RNase P-mediated inhibition of HIV-1 replication and pathogenesis [2001]

We may therefore hypothesise a correlation between HIV pathogenesis (and the main diagnostic questions related to it) being fully clarified at the end of the 1990s, and the number of Molecular biology RAs significantly decreasing in the corpus from 2001 onwards.

3.1.3 Immunology RATs

Immunology RATs typically convey research on two main topics, both related to the reaction triggered in the organism by HIV infection. On the one hand, they study various types of response to the virus (Chinen/Shearer 2002), and the impairment undergone by the immune system, charting a taxonomy of dysfunctions in its various arms, whereby lymphocytes are invaded by the virus so that the body, progressively depleted, becomes subject to opportunistic infections and tumours (Makvandi-Nejad 2018). This topic is evidenced in the corpus by examples such as the following:

(7) The virological and immunological consequences of structured treatment interruptions in chronic HIV-1 infection [2001]

---

6 In March 1985 the FDA licensed the first commercial blood test for HIV antibodies, followed in May 1992 by a 10-minute rapid test kit and in December 1994 by an oral test, the first to use fluids other than blood.

(8) HIV immunosuppression and antimalarial efficacy: Sulfadoxine-pyrimethamine for the treatment of uncomplicated malaria in HIV-infected adults in Siaya, Kenia [2006]

On the other hand, much of immunologists' agenda in the 1990s and 2000s was devoted to the experimental development and testing of candidate prophylactic (or therapeutic) HIV vaccines, which fact may be connected to the frequency of Immunology RATs in the corpus. As shown in Graph 1 (see Appendix), Immunology RATs are seldom published in the early years of the corpus (3.5% in 1991, 15% in 1996), but the trend is fully reversed in 2001, when they reach 35.4%, and later in the decade, with a peak in 2006 (343 RATs, 40.4%). Such a trend seems to be confirmed by the incidence, along the considered time span, of discourse about producing a vaccine against AIDS, as is testified in the corpus by the frequent occurrence of RATs along the lines of the following:

(9) Phase 1 safety and immunogenicity evaluation of a multiclase HIV-1 candidate vaccine delivered by a replication-defective recombinant adenovirus vector [2006]

(10) A randomized, partially blinded phase 2 trial of antiretroviral therapy, HIV-specific immunizations, and interleukin-2 cycles to promote efficient control of viral replication [2006]

The first mention of a vaccine being under study by Dr. Gallo at the NCI (supposedly to be produced within two years) occurred in April 1984; in 1993-94, AIDS being revealed as the leading cause of death in US patients aged 25-44, President Clinton established the National Office for AIDS policy at the White House, while Congress increased funding for research on candidate vaccines.[7] And while the 2004 G8 Summit advocated a Global HIV Vaccine Enterprise (i.e., a consortium of government and private groups whose task was to boost efforts; HIV.gov 2016), two clinical trials conducted in 2003 and 2007 resulted in failure to confirm vaccines' efficacy (Sekaly 2008). We

[7] In the 2000 State of the Union address, a Millennium Vaccine Initiative was launched as an incentive to further experimentation; 18 May 2001 was the first US HIV Vaccine Awareness Day.

may hypothesise said circumstances to be be linked to the conspicuous decrease in the publication of Immunology RATs in the corpus as of 2011, when their frequency drops by 15% (553 items, 27.8% of RATs), although this is mere speculation. What is known for sure is that, in 2006, Immunology RATs represent 28.6% of written scientific discourse on HIV (648 out of 1,725 items), partially to the advantage of Infectious diseases RATs.

3.1.4 Infectious diseases RATs

Infectious diseases RATs are generally concerned with both the systematic management of individual HIV/AIDS cases through highly active antiretroviral (HAART) protocols designed to prolong and enhance the life expectancy of patients and, on the other, with epidemiological surveillance on a global scale through population-based monitoring and evaluation programmes. As can be observed in Graph 1 (see Appendix), Infectious diseases RATs follow a similar trend to Immunology, in that both categories start being published extensively only as of 2001 onwards. While subsidiary in 1996 (19 items, 8.6%), Infectious diseases titles increase by 20% in 2001 (116 items, 28.7%), and again by approximately 12% in 2006 (355 items). Infectious diseases RATs maintain a publication frequency of over 30% in 2011 and 2016.

These data appear to be coherent with the rapid development, in the late 1990s and 2000s, of HIV/AIDS therapeutic protocols, and the spreading use of several classes of antiretroviral medications that, if taken in combination, are able to slow down the virus and stop the disease's progression.[8] In 1997, as HAART set in as the new standard of HIV care, a 47% decline in deaths in the US was reported for the first time since the onset of the epidemics (Stolberg 1997).

[8] The first successful antiretroviral drug (zidovudine AZT) became available in March 1987 (Hogg et. al 1998); in June 1994 the FDA approved the first HIV protease inhibitor (i.e., an agent preventing viral replication), which ushered in the era of Highly Active AntiRetroviral Therapy. Saquinavir was approved for prescription use in 1995, followed within four months by ritonavir and indinavir (Günthard *et al.* 2014).

Unsurprisingly, the topic of finding and fine-tuning a cure for AIDS is amply exemplified in the corpus by RATs such as the following:

(11) The effect of highly active antiretroviral therapy on cervical cytologic changes associated with oncogenic HPV among HIV-infected women [2001]

(12) Pre-Exposure Prophylaxis and Antiretroviral Resistance: HIV Prevention at a Cost? [2011]

The first disease ever to be debated at a UN General Assembly (1987), AIDS was ranked in 1999 by WHO as the fourth biggest cause of death worldwide (14 million casualties, 33 million HIV-positive people) and as the number one killer in Africa (HIV.gov 2016).[9] It thus seems coherent that Infectious diseases RATs in the corpus (especially as of 2001 onwards) should extensively deal with HIV epidemiology, especially in Africa and developing countries, as is testified by titles such as the following:

(13) Immunologic Criteria Are Poor Predictors of Virologic Outcome: Implications for HIV Treatment Monitoring in Resource-Limited Settings [2011]

(14) Age-disparate sex and HIV risk for young women from 2002 to 2012 in South Africa [2016]

If, by the end of 1985, at least one case of HIV/AIDS had been reported in all regions of the world (UNAIDS 2006), the 1990s and 2000s saw the joint effort of countries worldwide in order to follow a common surveillance framework within their own epidemiological settings. As a matter of fact, in the whole WoS corpus, Infectious diseases RATs, together with Immunology RATs, account for as many as 62% (3,136 items) of the total titles.

---

9 In 2002, UNAIDS reported that life expectancy in sub-Saharan Africa was reduced from 62 to 47 years as a result of AIDS (UNAIDS 2002); in 2013, Africa had ca. 6.8 million HIV diagnoses, which made it the region with the highest infection rate in the world (UNAIDS 2013).

3.2 Average sentence length

Table 2 (below) shows the average sentence length of RATs in the corpus, per year and per clinical specialty.

1986	
1. Public health	--
2. Molecular biology	--
3. Immunology	--
4. Infectious diseases	--
AVG 1986	--
1991	
1. Public health	12
2. Molecular biology	13.8
3. Immunology	12.5
4. Infectious diseases	10
AVG 1991	*12.1*
1996	
1. Public health	14.1
2. Molecular biology	15.8
3. Immunology	16.5
4. Infectious diseases	17.3
AVG 1996	*15.7*
2001	
1. Public health	15
2. Molecular biology	15.2
3. Immunology	15.2
4. Infectious diseases	15
AVG 2001	*15.1*
2006	
1. Public health	15.6
2. Molecular biology	15.6
3. Immunology	15.6
4. Infectious diseases	15.7
AVG 2006	*15.6*
2011	
1. Public health	17.2
2. Molecular biology	16.8
3. Immunology	19.3
4. Infectious diseases	19.7
AVG 2011	*18.3*

2016	
1. Public health	17.2
2. Molecular biology	17.1
3. Immunology	22.2
4. Infectious diseases	21.8
AVG 2016	*19.6*

Table 2. Average sentence length per year and per clinical specialty.

As can been seen in the Table, the average word number of RATs per year follows a steadily increasing trend, from 12.1 words (1991) to 15.6 (2006), 18.3 (2011) and 19.6 (2016). The average sentence length across clinical specialties appears – within the same year – to be quite uniform, with specialties substantially mirroring each other (e.g. 2001 and 2006), or in any case not showing dissimilar values (e.g. 1991, 1996, 2011), with the only possible exception of 2016, where Immunology and Infectious diseases RATs are 4-5 words longer than Public health and Molecular biology titles. As will be argued in Discussion, this seems coherent with the advancing status of medical research in Immunology and Infectious diseases, with respect to Public health and Molecular biology. As no regulations concerning title length are provided in the International Committee of Medical Journal Editors (ICMJE)'s *Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals*,[10] the increasing number of words can be interpreted as interdependent with the refinement of specialised knowledge about HIV.

3.3 Structural construction of RATs

On the basis of the above, RATs have been distinguished into four syntactic categories, according to different structural organizations of the material they provide, along a pragmatic continuum between information packaging and scientific attractiveness and persuasiveness (Consonni 2018; Sala/Consonni 2018; Wang/Bai 2007).

[10] Retrieved from <http://www.icmje.org/icmje-recommendations.pdf>.

Conclusive RATs (or declarative/full-sentence) are syntactically independent structures, hinging on finite verb forms which specify the semantic relationship among the lexical items in the sentence, as in the following examples:

(15) Household water insecurity is associated with a range of negative consequences among pregnant Kenyan women of mixed HIV status [2016 public health]

(16) CDC42 and Rac1 are implicated in the activation of the Nef-associated kinase and replication of HIV-1 [1996 molecular biology]

Interrogative RATs are structured in the form of questions, codifying meanings interrogatively rather than declaratively. For this reason, they can be used to either point to a possible gap in knowledge about certain phenomena, which the ensuing RA will deal with, or cast doubts over previous research on the same topic, etc. Interrogative RATs typically formulate "queries in need of reply, interpretation and conclusion" (Soler 2007: 100), as is shown by the following examples:

(17) Can data for programs for the prevention of mother-to-child transmission of HIV be used for HIV surveillance in Kenya? [2006 public health]

(18) Do type and duration of antiretroviral therapy attenuate liver fibrosis in HIV-hepatitis C virus-coinfected patients? [2006 immunology]

Nominal RATs are particularly condensed structures, either consisting of a single verbless expression, or containing non-finite verbal forms (such as gerund, participle, *to* + infinite, etc.). Typically occurring in "block language" (Straumann 1935), 'headlinese' (Garts/Berstein 1963) or economy grammar (Halliday 1967), they are an optimal resource for the codification of meanings in contexts with significant space constraints, such as newspaper headlines, book titles, advertising slogans, etc. Nominal constructions have a general preference for the omission of auxiliary verbs (*be*, *have*, *do*) and articles (*a/an*, *the*), and for passive voice and nominalization, as is shown by examples such as the following:

(19) Inpatient morbidity among HIV-infected male soldiers prior to their diagnosis of HIV-infection [1991 public health]

(20) Impaired development of HIV-1 gp160-specific CD8(+) cytotonic T cells by a delayed switch from Th1 to Th2 cytokine phenotype in mice with Helicobacter pylori infection [2001 immunology]

Compound RATs (or colonic/hanging, Hartley 2005b) are composed of two semantically interrelated parts – which can be declarative, interrogative and nominal phrases, clauses or full sentences – typically joined by a colon, full stop, dash or other punctuation mark (Hartley 2007; 553). With reference to their thematic structure, they form theme-rheme clusters, where the former part of the RAT generally introduces the article's topic, while the latter (usually an expanded noun phrase) details particular aspects of the topic that will be dealt with in the article, thus both pinpointing the title's relevance, by framing it in 'general'-'specific', 'cause'-'effect', 'problem'-'solution', 'question'-'answer' sequences, and highlighting the article's weight in terms of informative content. This can be observed in the following examples:

(21) Structure and stability of RNA/RNA kissing complex: Application to HIV dimerization initiation signal [2011 molecular biology]

(22) Viability and effectiveness of large-scale HIV treatment initiatives in sub-Saharan Africa: Experience from Western Kenya [2006 immunology]

Table 3 shows the distribution of syntactic constructions per year and per clinical specialty.

	Conclusive	*Interrogative*	*Nominal*	*Compound*
1986				
1. Public health	0	0	0	0
2. Molecular biology	0	0	0	0
3. Immunology	0	0	0	0
4. Infectious diseases	0	0	0	0
SUBT. 1986: 1 (100%)	*0 (0%)*	*0 (0%)*	*0 (0%)*	*0 (0%)*
1991				

1. Public health	0	0	10	8
2. Molecular biology	1	0	11	3
3. Immunology	1	0	1	0
4. Infectious diseases	0	0	1	0
SUBT. 1991: 36 (100%)	*2 (5.6%)*	*0 (0%)*	*23 (63.9%)*	*11 (30.5%)*
1996				
1. Public health	0	1	8	4
2. Molecular biology	20	1	25	18
3. Immunology	6	1	24	2
4. Infectious diseases	3	1	13	2
SUBT. 1996: 129 100%)	*29 (22.5%)*	*4 (3.1%)*	*70 (54.2%)*	*26 (20.2%)*
2001				
1. Public health	0	2	15	13
2. Molecular biology	24	0	21	13
3. Immunology	34	1	84	24
4. Infectious diseases	16	0	65	35
SUBT. 2001: 347 100%)	*74 (21.3%)*	*3 (0.9%)*	*185 (53.3%)*	*85 (24.5%)*
2006				
1. Public health	0	4	17	25
2. Molecular biology	26	0	39	8
3. Immunology	65	1	197	80
4. Infectious diseases	58	3	201	93
SUBT. 2006: 817 100%)	*149 (18.2%)*	*8 (1%)*	*454 (55.6%)*	*206 (25.2%)*
2011				
1. Public health	0	6	88	62
2. Molecular biology	18	0	59	51
3. Immunology	102	7	240	204
4. Infectious diseases	88	12	346	167
SUBT. 2011: 1,450 (100%)	*208 (14.3%)*	*25 (1.7%)*	*733 (50.6%)*	*484 (33.4%)*
2016				
1. Public health	1	4	109	102
2. Molecular biology	21	3	50	32
3. Immunology	44	17	298	289
4. Infectious diseases	160	28	395	172
SUBT. 2016: 1,725 (100%)	*226 (13.1%)*	*52 (3%)*	*852 (49.4%)*	*595 (34.5%)*
TOTAL 1986-2006: 4,504 (100%)	688 (15.3%)	92 (3%)	2,317 (51.4%)	1,407 (31.2%)

Table 3. Distribution of syntactic constructions per year and per clinical specialty.

As can be seen in Table 3, the frequency of *conclusive* RATs in the WoS corpus gradually but steadily decreases from 1996 onwards: they go from 22.5% in 1996 (29 items) to 13.1% in 2016 (226 items). This seems to indicate that, as research on HIV develops in the late 1990s and 2000s, conclusive titles no longer appear to be a common choice for the structuring of RATs. On the one hand, full-sentence titles may in fact be associated to pragmatic necessities of scientific discourse such as the straightforward presentation of informative material and/or informative density/attractiveness, thus matching the author's need to rapidly inform readers and fellow experts about the contents of the article, while readers in turn need to "know as early as possible in the reading process whether or not the paper contains anything" of relevance (Haggan 2004: 296).

On the other hand, *conclusive* sentences may sound as confident assertions presenting hypotheses as "statements of fact", generally in the present tense, reproducing what has been tagged as the "block language" of newspaper headlines (Quirk/Greenbaum 1973; Straumann 1935). Furthermore, 100% of conclusive RATs in the WoS corpus are in the present tense, which may indicate an optimistic epistemological attitude on the part of the researcher, who may cast on his/her text the idea that what s/he is reporting "stands true for all time or is not simply a one-off occurrence", or the belief that "the method, measurements, calculation etc. employed have yielded impregnable results (Haggan 2004: 297). Significantly, only 4 out of 688 conclusive titles in the corpus (1 in 1991 Molecular biology, 3 in 2006 Immunology) are accompanied by the use of hedges (Hyland 2005), and specifically by the modal verb *can*, while there is no use of such typical markers of reduced epistemicity in scientific discourse as the other central modals (for instance, *may*, *might* and *could*), semi-modal verbs (such as *ought to*, *need to*) or lexical-modal auxiliaries or idioms, such as *be due*, *bound*, *certain*, *likely* etc. (Hyland 1998).

As regards *interrogative* RATs, assuming them to be syntactic expressions of doubt, contradicting to some extent the status of scientific research as the formulation of exact answers to complex questions, it seems coherent that (as can be seen in Table 3) they are the least preferred strategy for codifying meanings in the corpus, totalling 92 occurrences (3%) from 1986 to 2016.

As can be seen in Table 3, *nominal* RATs are by far the preferred strategy across all clinical specialties, largely outnumbering all other constructions in all considered years and totalling 51.4% of structures (2,317 occurrences). A closer look at the data however reveals that, albeit the dominant category, nominal titles tend to follow a mildly decreasing trend, from 63.9% (1991) to 49.4% (2016), in correlation with an increasing trend of the last syntactic category, compound structures.

Compound RATs are the second preferred resource for title structuring in the corpus (see Table 3), with a regular tendency to increase over time, from 20.2% to 34.5% (2016). By looking at the relative incidence of nominal *vs*. compound syntax along the considered time span, it can moreover be noticed that the difference in percentage between the two categories tends to decrease, from 63.9% nominal *vs*. 30.5% compound (1991) to 49.4% nominal *vs*. 34.5% compound in 2016.

3.3.1 Structural constructions within and across specialties

Graphs 2-6 (see Appendix), calculated and drawn on the basis of Table 3, show the distribution of syntactic constructions in, respectively, the whole corpus, comprising RATs in all clinical specialties (Graph 2), and in Public health, Molecular biology, Immunology and Infectious diseases RATs (Graphs 3-6).

As evidenced in Graph 2 (see Appendix), by comparing syntactic constructions across specialties it appears that all four research areas privilege the use of nominal RATs, which range between 63.9% (1991) and 49.4% (2016). However, the proportion between nominal and compound structures eventually tends to become more evenly balanced, especially in 2011 and 2016, where an increasing preference for compound syntax can be observed, and the difference between the two categories decreases from ca. 50% to ca. 15%. It can also be noticed that interrogative structures are seldom used in the corpus, ranging between 0% and 3.1%, while the third preferred resource for titling constructions, conclusive syntax, has a tendency to decrease over time (22.5% in 1996 *vs*. 13.1% in 2016) to

the advantage of compound syntax. In fact, the 15% increase in compound titles from 1996 to 2016 does not appear to be motivated by a significant decrease in nominal structures (going down by ca. 5% between 1996 and 2016), but by a 10% decrease in conclusive titles (21.3% in 1996 *vs*. 13.1% in 2016). By observing Graphs 3-6 (see Appendix) we can further compare the distribution of constructions within each specialty.

Never a viable choice in Public health, conclusive titles are the third preferred choice in Infectious diseases and Immunology RAs (with the only possible exception of 1991 Immunology), while only Molecular biology shows a preference for this type of construction, which, especially in 1996 (31.2%) and 2006 (35.6%), appears to be one of two preferred strategies (along with nominal titles). (We can also notice a peak in 2001 Molecular biology, where conclusive titles are the number one resource, i.e., 41.4%). Although they are the least frequent constructions in all categories, interrogative titles are the number three strategy in Public health RAs, ranging between 0% and 8.7% (2006), at the expense of conclusive titles, which are almost never used (0%-0,4%). As concerns the distribution of nominal *vs*. compound structures, the former by large appear as the preferred choice in Infectious diseases RAs, where they range from 33.9% (1991) to 49.4% (2016), and in Molecular biology, where they appear in 36.2%-73.3% of titles. But the trend is significantly different in Public health RAs, where nominal titles range between 37% and 61.5%, and – quite similarly – compound titles range between 30.8% and 54.3%, peaking in 2006, when they outnumber nominal titles (54.3% *vs*. 37%). In Immunology RATs, nominal structures generally prevail over compound structures, with the only exception of 2016, when the proportion is almost 1:1 (44.6% *vs*. 46%).

To sum up, it is possible to claim that the distribution of syntactic constructions within each specialty is not as homogeneous as it may appear from Graph 2. Graphs 3-6 show that, as we move from 1986 to 2016, different strategies are chosen within each specialty in order to structure meanings about HIV. Public health RATs tend to prefer nominal and compound constructions, never recurring to declarative syntax; Molecular biology RATs tend to privilege nominal and (to a lesser extent) compound and conclusive constructions,

excluding interrogative structures; Immunology tends to privilege nominal structures, with a tendency, over time, to use more compounds and less declaratives; Infectious diseases RATs tend to use nominal structures much more frequently than the other categories, although with, in turn, a tendency to use more compounds and less conclusive structures as time goes by. Such preference for compound structures seems coherent with Immunology and Infectious diseases RATs having the highest average sentence length (see section 3.2).

3.4 Textualization of EBM methodology in RATs

Table 4 provides data on the frequency of a significant phenomenon deployed in the textualization strategies within syntactic constructions – and especially in compound RATs – in order to package and sequence information for readers. In particular, it shows the distribution, across titling constructions, of expanded nominal phrases (NPs) bearing lexical reference to Evidence Based Medicine study designs. Introduced in the early 1990s and officially codified later in the decade (Sackett *et al*. 1996), EBM is based upon a 'pyramid of evidence' (see Figure 1), where seven ascending levels in experimental methodology – or study designs – are ranked according to increasing result reliability in statistical terms and decreasing (or minimized) bias from confounding variables (Greenhalgh 2010).

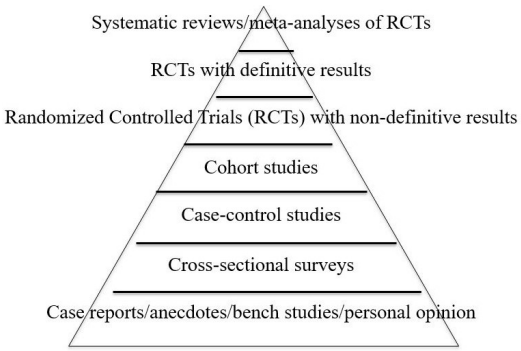


Figure 1. The EBM pyramid of evidence

Each of the seven levels represents the increasing epistemicity, in terms of the provided clinical evidence, of a study design (i.e., the methodology employed in research, as presented in the Methods section of the RA, in terms of participants, implements and procedures). The pinnacle is represented by the robust, generalizable evidence inherent to systematic reviews and meta-analyses of Randomized Controlled Trials, while at the bottom lies the anecdotal knowledge produced by case reports.

	Conclusive	*Interrogative*	*Nominal*	*Compound*	*Subt.*
1986	0	0	0	0	0 (0%)
1991	0	0	1	3	4 (11%)
1996	0	0	3	8	11 (8.5%)
2001	0	0	4	18	22 (6.3%)
2006	0	0	24	100	124 (15.2%)
2011	0	0	184	320	504 (34.6%)
2016	0	0	201	411	612 (35.5%)
TOT. (1,277)	0 (0%)	0 (0%)	417 (32.6%)	860 (76%)	

Table 4. Distribution of expanded NPs indicating EBM study designs per year and per syntactic construction.

As shown in Table 4, in the WoS corpus nominal and compound structures particularly tend to rely on expanded NPs lexicalizing EBM study designs, as in the following examples (emphasis added):

(23) *Randomized*, *double-blind* comparison of two nelfinavir doses plus nucleosides in HIV-infected patients (Agouron study 511) [2001 immunology, nominal]

(24) The prevalence and correlates of HIV and undiagnosed infection among men who have sex with men in Hanoi, Vietnam: Findings from a *cross-sectional, biobehavioral study* [2016 public health, compound]

The frequency of EBM-related NPs steadily increases over time, from 11% (4 out of 36 items in 1991) to 15.2% (2006), 34.6% (2011) and 35.5% (2016). Of such NPs, 32.6% appear in nominal constructions (417 occurrences), while 67.4% in compounds (860 occurrences).

	Public health	*Molecular biology*	*Immunology*	*Infectious diseases*
1986	0	0	0	0
1991	3	0	0	0
1996	1	0	3	4
2001	3	0	10	5
2006	5	8	42	45
2011	40	27	113	140
2016	71	47	125	168

Table 5. Distribution of EBM-related NPs in compound structures per year and per clinical specialty

Table 5 (above) and Chart 1 (see Appendix) provide further data as to the specific frequency of such occurrences in compound constructions per year and per clinical specialty. As can be seen, the specialties that most showcase EBM methodology in compound constructions are Immunology and Infectious diseases, particularly from 2006 onwards. This seems coherent with the data on clinical coverage presented in sections 3.1.3 and 3.1.4.

Table 4 furthermore evidences that the most productive category with respect to the textual codification of EBM methodology is compound syntax. These data appear to be in line with the textual affordances offered by compound titling, in terms of information patterning strategies. While the thematic part of compound RATs mainly focusses on the article's topic, the extended NP following punctuation and occupying the rhematic position generally provides specifics as to *how* the topic is addressed in the article.

As of the mid-1990s onwards, as Table 4 shows, the rheme/filler slot tends to be occupied by NPs indicating the research's study design, thus immediately elucidating the article's ranking, in terms of scientific prestige, within the EBM hierarchy of knowledge. In such NPs, the lexicalisation of study design terminology functions as a cognitive cue to the reader, who is invited to assess the study's methodology by reading the appropriate IMRD section within the article. Moreover, the sequential theme/rheme, gap/filler information patterning – which textualizes the positioning of a piece of research within evidential knowledge – tends to coincide with the structure's patterning in terms of Information Unit. While the thematic/given part of the compound generally presents a specific clinical aspect of

HIV/AIDS, the rhematic/new part of the structure/message tends to frame the research's ranking within the EBM paradigm. This seems compatible with both the increasing sentence length presented in section 3.2, and the tendency, analysed in section 3.3.1, of all clinical specialties (with the possible exception of Molecular biology) to significantly increase their use of compound syntax in the course of the 2000s, particularly at the expense of conclusive titles.

4. Discussion and conclusions

In 1986, as the name HIV entered the lexicon of physicians worldwide, scant knowledge was available on the cause, transmission, evolution and possible treatment of this pathology. Fatalism, scandal and marginalization were the first reactions of US industrialised society vis-à-vis a global epidemic. In the following years, research established a completely new field of knowledge, and a new type of discourse was implemented within specialised communication. From a discourse-analytical angle, this study aims at putting forward a functional and epistemological correlation between the disciplinary construction of specialised medical knowledge about HIV and the linguistic evolution and variation of expert-to-expert communication conveying such knowledge. Because of the semiotic affordances they offer (i.e., informativity, attractiveness, and the indexing of a research article's topic and methodological positioning), and the multifaceted pragmatic functions (i.e., argumentative, persuasive and metadiscursive) they serve, RATs are here assumed as a key genre for investigating both the disciplinary matrix and the communicative features of medical discourse about HIV.

From 1986 to 2016, the interaction among the scientific scopes and purposes of Public health, Molecular biology, Immunology and Infectious diseases has covered a whole cline of meanings in pathology, a spectrum of conceptual areas with which HIV/AIDS has progressively been associated in its history, and which are stratified in its contemporary perception. In the late 1980s, HIV was first framed

by public health studies within the urgent framework of global pandemics and contagion alarm, in concert with political and news media stigmatizing discourse with respect to certain social categories first evidenced as being exposed to infection. In the following decade, HIV became the object of closer biological investigation, which, by bringing into focus the issues of its pathogenesis and diagnosis, shifted both experts' and the lay public's attention from social stigma towards the actual complexity of multiple aetiological factors. As of the late 1990s onwards – significantly, at one time with the standardisation of EBM as a shield against bias – immunology and infectious diseases became the dominant perspectives in HIV research, covering two thirds of specialised written discourse worldwide. A progressive amount of increasingly detailed and substantial knowledge about the disease and its biological and social impact, as well as about its potential control through candidate vaccines and antiretroviral treatment, thus became available, pointing – in the 2000s – in the direction of its global systematic monitoring, containment and treatment through massive statistics-based surveillance and therapy. Epidemics and contagion, pathogenesis and diagnosis, damage and immunity, therapy and epidemiology thus broadly seem to be the dichotomies structuring HIV research in the US from 1986 to 2016.

In the light of the above, sections 3.2 and 3.3 have analysed a corpus of 4,504 RATs, in order to quantify and interpret the structural and textual resources employed in regard to the codification and dissemination of both the disciplinary matrix of HIV discourse (research question 1) and the inception of Evidence Based Medicine (research question 2). Results can now be discussed with reference to both questions. As concerns the average title length (see section 3.2 and Table 2), the brevity of RATs in the earlier years of the corpus may stem from the communicative urgency of early research, when hypotheses and results needed to be quickly packaged and spread in "brief and succinct" form (Haggan 2004: 294). Conversely, the increasing length of structures over time may indicate a growth in the informative content of RAs, and, more specifically, an increase in the number of lexical items they contain. This occurrence may mirror a common 'time factor' trend in the length of scientific titles (Yitzhaki

1997), and may be interpreted as a twofold phenomenon. On the one hand, the more keywords a title contains, the more accurately it can showcase its results, and the more chances are that it will be retrieved from database or online queries, thus broadening the range of its diffusion, especially in the digital era.

On the other hand, the evolution of HIV research across different specialties may also have triggered an increase in RAT length, for as a field of knowledge becomes more articulate and reified, titles can reasonably be expected to become more complex, in order to mirror "the development, refinement and extension both of underlying theories and of more and more complex research methods and procedures" (White/Hernandez 1991: 731). This is particularly likely to happen in the hard sciences, traditionally having more informative titles than the popular sciences (Hjørland/Nielsen 2001), whereby increasing knowledge specialisation calls for "more words to express a given piece of research" (White/Hernandez 1991: 731), so that longer titles may function as vehicles to disseminate new achievements, also providing additional room for the presentation of a given piece of research's positioning within the EBM methodological framework.

In parallel, the patterns and variations among the four syntactic categories of RATs, presented in section 3.3 (see Table 3, and Graphs 2-6 in the Appendix), also evidence the functional and epistemological interdependence between specialised knowledge on HIV and the linguistic configuration of expert-to-expert written medical discourse. Such interconnection can be observed at various levels. The different frequency patterns of conclusive constructions among clinical specialties may suggest a conflation in RATs between scientific and promotional language, whereby 'headlinese' effects can be employed to encode some degree of epistemological certainty on the topic – especially in the case of Molecular biology RATs, which tend to promote the use of such constructions. In the WoS corpus, conclusive titles may thus be read as a typical strategy for the kick-off stage of HIV research, as the very complexity of the disease's biology and epidemiology was about to be taken into consideration. This hypothesis seems confirmed by a substantial decline in declarative syntax over time (see Graph 2), and may also apply to the low

frequency of interrogative structures in the corpus, for little use of questions may reflect lesser need for the structural expression of scientific dilemma, especially as research developed in the 1980s and 1990s.

The high incidence of nominal structures across all specialties seems to deserve a more articulated interpretation. With their high capacity for highlighting a discipline's theoretical and experimental concerns, as well as its keywords, nominal constructions may be associated with science's constant need for maximally condensed lexicalisation strategies, whereby increasingly complex research on increasingly subtle topics calls for effective linguistic strategies for both encoding and emphasizing such complexity in a limited amount of space and time. The overall prevalence of nominal structures may thus be traced to the prototypical classificatory nature of biomedical discourse, which tends to handle its subject matter in taxonomical fashion (Soler 2007: 101), evidencing layer after layer of experimental knowledge through the piling up of lexical items in the construction of extended noun phrases. This view is also confirmed by Haggan (2004: 307), according to whom a NP, accompanied by one (or more than one) post-modifying prepositional phrases and/or elaborate pre-modification, is the most viable strategy for traditional scientific title-patterning, in that it provides RATs with both informative precision and explicitness as to their contents (especially through the use of post-modifiers) and block-language attractiveness, as attained by the effective use of shorter, and generally more evenly balanced, sets of pre-modifiers (Rush 1988).

The increasing frequency of compound titling structures in the corpus, noticeable within and across all specialties, especially as of 2001 onwards, seems in turn to mirror – through its typical theme/rheme, gap/filler, topic/method patterning – science's growing need for complementing traditional scientific informativity with attractiveness and persuasiveness (Haggan 2004; Hartley 2007; Bhatia 2004). Compound titles, as claimed by Hartley (2007: 558), both attract and inform readers by means of their information sequencing: the audience's curiosity is elicited by the thematic part of the cluster, which works as a gap (usually introducing a research question), while the filler slot presents insights as to possible answers and/or as to how

the question will be dealt with (Haggan 2004: 302). In opposition to te traditional nominal structures, which present results synoptically (much in the fashion of conceptual maps to the article's findings), and immediately enable readers to grasp the essential input of the article, compound syntax works along sequential "add-on" theme/rheme protocols, pivoting on the opposite principle, that is, the presumption of the audience's ignorance (Haggan 2004: 203).

Such conflation of pragmatic purposes may suggest further research into medical RATs as a hybrid genre between scientific and promotional language, and possibly as an advanced textual resource within the hybridity traditionally inherent to the use of conclusive ('headlinese') RATs. The markedly standardised, hierarchic formulations of traditional medicine, as typified by nominal structures, meant for the selective skimming of information on the part of researchers before lay dissemination, are complemented in compound titles by more articulated and horizontal functional protocols, placing emphasis on argumentative, persuasive and metadiscursive functions. In the WoS corpus, the frequency of compound titles appears to significantly increase over the 2000s (see Graph 2 in the Appendix): this marks a watershed in the pragmatic and epistemological scope of specialised medical communication, from the gate-keeping, esoteric exchange of individual expertise (which, within the EBM hierarchy of knowledge, would rank quite low) to the peer-to-peer sharing of evidence-based knowledge. In their informative and persuasive functions, compound RATs seem to textualize scientific discourse as an arbitration process between hypotheses and phenomena, and as the social negotiation of methodologies and expectations (Kuhn 1962); they can also be interpreted as markers of a scientist's self-aware, discipline-mediated and socially negotiated positioning within the boundaries of a scientific system.

Like nominal titles, compound titles also contain a high number of lexical items, which boost the visibility of research in scientific databases (and online searches) and offer additional room for showcasing its experimental advancements and methodological reliability. No longer dependent from the full RA, a compound title is *per se* a semantically autonomous structure, performing key cognitive functions, i.e., facilitating the decoding of the article, directing

"attentional focusing while reading", providing "text information by establishing relations between different elements", determining "the relative importance of information supplied in a text" and building readers' "cognitive representation" (Eyrolle / Virbel / Lemarié 2008: 242). But, as shown in section 3.4, the textualization strategy deployed in compound syntax appears to be more audience-oriented – in regard to today's Web-based and EBM-centred scientific paradigm – than the piling up of pre- and post-modifiers. The chain of knowledge transmission in compound syntax is sequential and horizontal, rather than synoptic and vertical: communication is centred on the reader's two-step appraisal (curiosity followed by cognition) rather than on the writer's authoritative configuration of contents. In the WoS corpus, such information packaging system increases in frequency as of 2001 onwards – in all clinical specialties – in concert with the textualization of EBM methodology (see Table 5, and Chart 1 in the Appendix), whereby a study's ranking within the pyramid of evidence is showcased in the rhematic part of the cluster. Such resource may be considered as a metadiscursive marker of evidentiality (Hyland 2005), i.e., a textual device indicating "the source of speakers' knowledge" (Johnstone 2009: 30) by attributing "information or opinion in a text to sources which may be animate or inanimate" (Hunston 2003: 181), such as evidence provided by empirical research (lab experiments, trials, etc). Embedded in the increasing use of such strategy is the growing need of scientists to engage in (and comply with) discourse legitimization issues in the EBM paradigm. Beside their typical pragmatic functions (i.e., informativity, attractiveness and retrievability of articles), RATs thus seem, as a genre, to have developed a crucial epistemological function.

To conclude, the phenomena quantified and interpreted in this study overall indicate that, along with research on HIV, the language of specialised written medical communication has undergone significant changes in the last three decades. In the context (and in the face) of a life-threatening global epidemic, and of the scientific efforts undertaken by institutions in industrialised Western countries – among which the US is the number one funder and beneficiary, as well as the number one discourse generator – the traditionally conservative language of medicine has shown signs of major changes within its

syntactic and textual features. The trends identified point in the direction of changing functional conventions and shifting epistemological attitudes towards the production, validation and dissemination of medical knowledge. More than at producing conclusive results, however, this paper has aimed at suggesting that linguistic and communicative changes within the specialised knowledge system of HIV discourse should be investigated against the complex backdrop of the multi-disciplinary clinical paradigm established in order to identify, define, diagnose and combat this pathology, as well as of the worldwide implementation of Evidence Based methodology in the last decades.

References

Anthony, L. 2016. *AntConc* (Version 3.4.4) [Computer Software]. Tokyo: Waseda University.

Bazerman, Charles 1988. *Shaping Written Knowledge. The Genre and Activity of the Experimental Article in Science*. Wisconsin: University of Wisconsin Press.

Bhatia, Vijay K. 2004. *Worlds of Written Discourse. A Genre-Based View*. London: Continuum.

Bucchi, Massimiano 1998. *Science and the Media. Alternative Routes in Scientific Communication*. New York: Routledge.

Calsamiglia, Helena 2003. Popularization Discourse. *Discourse Studies* 5/2, 139-146.

Calsamiglia, Helena / Dijk, Teun A. van 2004. Popularization Discourse and Knowledge about the Genome. *Discourse & Society* 15/4, 369-389.

CDC (Center for Disease Control and Prevention) 1987. Recommendations for Prevention of HIV Transmission in Health-Care Settings. *JAMA* 258/10, 1293-1305.

CDC (Center for Disease Control and Prevention) 2017. What Is Public Health? Retrieved from <https://www. cdcfoundation.org/what-public-health>.

Chinen, Javier / Shearer, William T. 2002. Molecular Virology and Immunology of HIV Infection. *Journal of Allergy and Clinical Immunology* 110/2, 189-198.

Consonni, Stefania 2018. HIV Discourse in the *British Medical Journal*, 1985-2005: The Impact of Digital Literacy and Evidence-Based Medicine on Syntactical Patterns and Variations in RA Titles. In Baldry, Anthony / Bianchi, Francesca / Loiacono, Anna (eds) *Representing and Redefining Specialised Knowledge: Medical Discourse*, Bari: ESE-Salento University Publishing, 73-96.

Eyrolle, Hélène / Virbel, Jacques / Lemarié, Julie 2008. Impact of Incomplete Correspondence between Document Titles and Texts on User's Representations: A Cognitive and Linguistic Analysis Based on 25 Technical Documents. *Applied Ergonomics* 39, 241-246.

Fortanet, Inmaculada *et el*. 1998. Linguistic Analysis of Research Article Titles: Disciplinary Variations. In Vázquez, Ignacio / Guillén, Ignacio (eds) *Perspectivas Praguietices en Linguistica Aplicada*. Zaragoza: Anubar, 443-447.

Frieden, Thomas R. *et al*. 2005. Applying Public Health Principles to the HIV Epidemic. *New England Journal of Medicine* 353: 2397-2402. doi: 10.1056/NEJMsb053133

Garzone, Giuliana 2006. *Perspectives on ESP and Popularization*. Milano: CUEM.

Garzone, Giuliana / Catenaccio, Paola 2008. *Language and Bias in Specialized Discourse*. Milano: CUEM.

Genette, Gérard 1987. *Seuils*. Transl. *Paratexts: Thresholds of Interpretation*. Cambridge: Cambridge University Press, 1997.

Gotti, Maurizio 2013, The Analysis of Popularization Discourse: Conceptual Changes and Methodological Evolutions. In Kermas, Susan / Christiansen, Thomas (eds) *The Popularization of Specialized Discourse and Knowledge across Communities and Cultures*. Bari: Edipuglia, 9-32.

Günthard, Huldrych *et al.* 2014. Antiretroviral Treatment of Adult HIV Infection: 2014 Recommendations of the International Antiviral Society-USA Panel. *JAMA* 312/4, 410-425. doi:10.1001/jama.2014.8722

Greenhalgh, Tricia 2010. *How to Read a Paper. The Basics of Evidence Based Medicine*, Oxford: Wiley/Blackwell.

Haggan, Madeline 2004. Research Paper Titles in Literature, Linguistics and Science: Dimensions of Attraction. *Journal of Pragmatics* 36, 293-317.

Halliday, Michael A.K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.

Hartley, James 2005a. Improving that Title: The Effects of Colons. *European Science Editing* 31, 45-47.

Hartley, James 2005b. To Attract or to Inform: What Are Titles for? *Journal of Technical Writing and Communication* 35/2, 203-213.

Hartley, James 2007, Planning that Title: Practices and Preferences for Titles with Colons in Academic Articles. *Library & Information Science Research* 29, 553-568.

HIV.gov 2016. A Timeline of HIV/AIDS. Retrieved from <https://www.hiv.gov/sites/default/files/aidsgov-timeline.pdf>.

Hogg, Robert S. *et al.* 1998. Improved Survival among HIV-Infelcets individualas following initiation of antiretroviral therapy. *JAMA* 279/6, 450-454. doi:10.1001/jama.279.6.450

Hyland, Ken 1998. *Hedging in Scientific Research Articles*. Amsterdam: Benjamins.

Hyland, Ken 2005. *Metadiscourse. Exploring Interaction in Writing*. London: Continuum.

Hjørland, Birger / Nielsen, Lykke K. 2001. Subject Access Points in Electronic Retrieval. *Annual Review of Information Science and Technology* 35, 249-298.

Hunston, Susan 2003. Evaluation and the Planes of Discourse: Status and Value in Persuasive Texts. In Hunston, Susan / Thompson, Geoff (eds) *Evalutation in Texts: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 176-207.

Jaime Sisó, Mercedes 2009. Anticipating Conclusions in Biomedical Research Article Titles as a Persuasive Journalistic Strategy to Attract Busy Readers. *Miscelánea* 39, 29-51.

Johnstone, Barbara 2009. Stance, Style, and the Linguistic Individual. In Jaffe, Alexandra (ed) *Stance: Sociolonguistic Perspectives*. Oxford: Oxford University Press, 29-62.

Kozminsky, Ely 1977. Altering Comprehension: The Effect of Biasing Titles on Text Comprehension. *Journal of Applied Research in Memory and Cognition* 5/4, 482-490.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Makvandi-Nejad, Shokouh (2018). Human Immunodeficiency Virus (HIV). British Society for Immunology. Retrieved from <https://www.immunology.org/file/2001/download?token=sy-0r6ZB>.

Martin, Shannon 1998. How News Gets from Paper to Its Online Counterpart. *Newspaper Research Journal* 19/2, 64-73.

Myers, Greg 2003. Discourse Studies of Scientific Popularization: Questioning the Boundaries. *Discourse Studies* 3, 265-279.

Quirk, Randolph / Greenbaum, Sidney 1973. *A University Grammar of English*. London: Longman.

Raffo, Mariana 2016. Translation and Popularization: Medical Research in the Communicative Continuum. *Meta* 61, 163-175.

Rush, Susan 1998. The Noun Phrase in Advertising English. *Journal of Pragmatics* 29, 155-171.

Sala, Michele / Consonni, Stefania 2018. Titles in Medicine and Science Popularization. In Bondi, Marina / Cacchiani, Silvia / Cavalieri, Silvia (eds) *Knowledge Dissemination at a Crossroads: Genres and New Media Today*. Newcastle Upon Tyne: Cambridge Scholars Publishing, 16-32.

Scott, Mike 2017. *WordSmith Tools version 7*. Stroud: Lexical Analysis Software.

Sekali, Rafik-Pierre 2008. The Failed Merck HIV Vaccine Study: A Step Back or a Launching Point for Future Vaccine Development? *Journal of Experimental Medicine* 205/1, 7-12.

Shinn, Terry / Whitley, Richard (Eds.) 1985. *Expository Science. Forms and Functions of Popularisation*. Dordrecht: Reidel.

Smith, Richard 2000. Informative Titles in the BMJ. *British Medical Journal* 320, 915.

Soler, Viviana 2007. Writing Titles in Science: An Exploratory Study. *English for Specific Purposes* 26, 90-102.

Stolberg, Sheryl G. (1997). US Says AIDS Cases Fell in '96, Ending 16-Year Rise. *The New York Times* September 19, A00026. Retrevied from <https://www.nytimes.com/1997/09/19/us/us-says-aids-cases-fell-in-96-ending-16-year-rise.html>.

Straumann, Heinrich 1935. *Newspaper Headlines. A Study in Linguistic Method*. London: Allen & Unwin.

UNAIDS 2002. Report on the global HIV/AIDS epidemic. Retrieved from <http://data.unaids.org/pub/report/2002/brglobal_aids_report_n_pdf_red_en.pdf>.

UNAIDS 2006. 25 Years of AIDS. Retrieved from <http://www.uncares.org/common/docs/25YearsofAIDS_en.pdf>.

UNAIDS 2013. Report on the global AIDS epidemic. Retrieved from <http://files.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf>.

UNAIDS 2018. 2017 Global HIV Statistics. Retrieved from <http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf>.

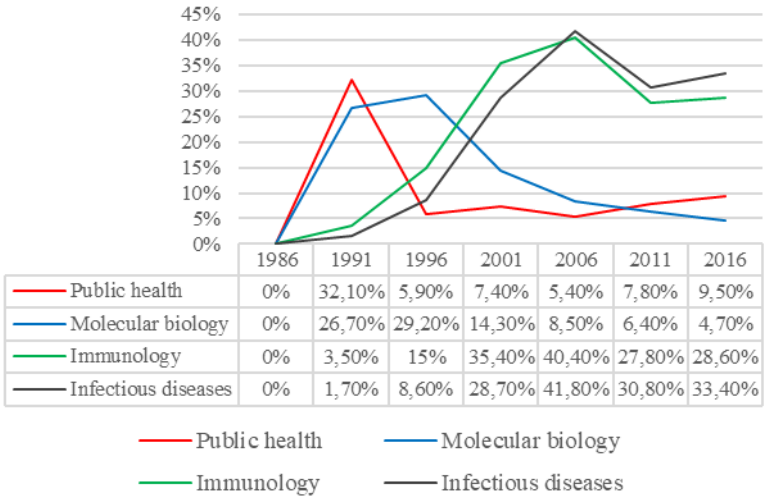
WHO (World Health Organization) 2018. HIV surveillance, estimates, monitoring and evaluation. Retrieved from < http://www.who.int/hiv/topics/me/en/>.

Yitzhaki, Moshe 1997. Variations in Informativity of Titles of Research Papers in Selected Humanities Journals: Comparative Study. *Scientometrics* 38/2, 219-229.

Wang, Yan / Bai, Yongquan 2007. A Corpus-Based Syntactic Study of Medical Research Article Titles. *System* 35, 388-399.

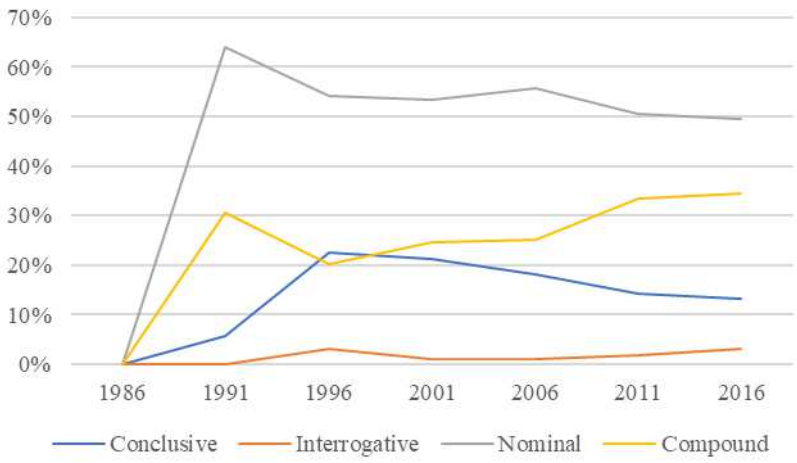
White, Arden / Hernandez, Nelda R. 1991. Increasing Field Complexity Revealed through Article Title Analyses. *Journal of the American Society for Information Science* 42/10, 731-734.

APPENDIX

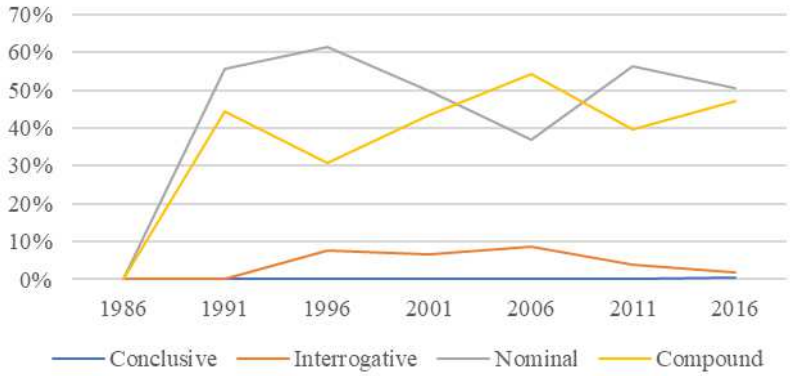


	1986	1991	1996	2001	2006	2011	2016
Public health	0%	32,10%	5,90%	7,40%	5,40%	7,80%	9,50%
Molecular biology	0%	26,70%	29,20%	14,30%	8,50%	6,40%	4,70%
Immunology	0%	3,50%	15%	35,40%	40,40%	27,80%	28,60%
Infectious diseases	0%	1,70%	8,60%	28,70%	41,80%	30,80%	33,40%

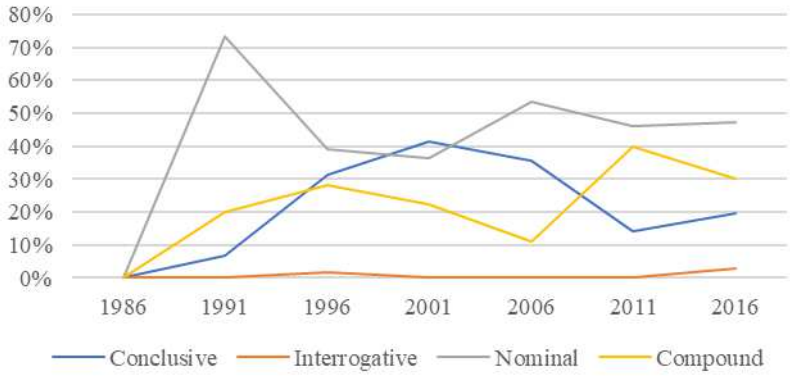
Graph. 1. No. of RATs per year and per clinical specialty.



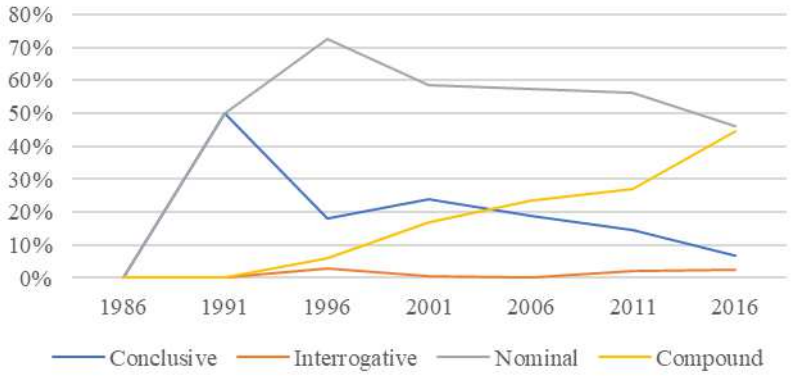
Graph 2. Distribution of syntactic constructions per year (all clinical specialties).



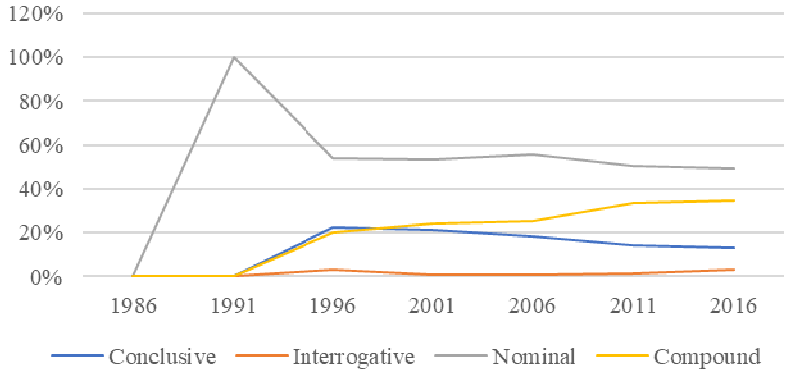
Graph 3. Distribution of syntactic constructions in Public health RATs.



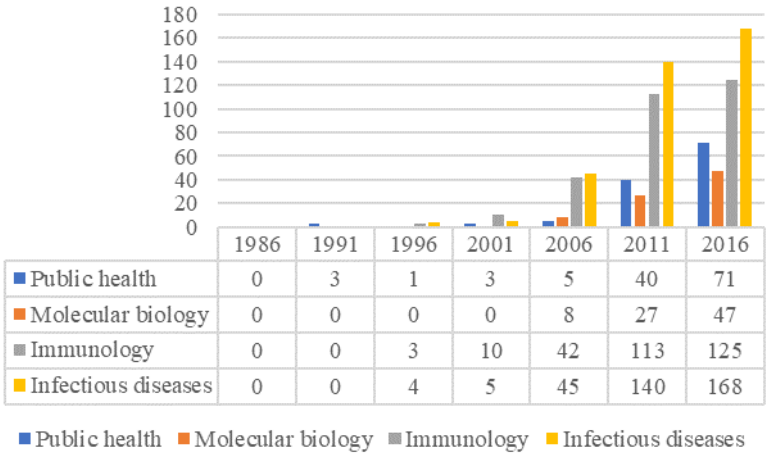
Graph 4. Distribution of syntactic constructions in Molecular biology RATs.



Graph 5. Distribution of syntactic constructions in Immunology RATs.



Graph 6. Distribution of syntactic constructions in Infectious diseases RATs.



	1986	1991	1996	2001	2006	2011	2016
Public health	0	3	1	3	5	40	71
Molecular biology	0	0	0	0	8	27	47
Immunology	0	0	3	10	42	113	125
Infectious diseases	0	0	4	5	45	140	168

Chart 1. Distribution of EBM-related NPs in compound structures per clinical specialty.