新語 INTERNET の脱大文字化：
the Internet か the internet か

久　屋　愛　実

1．はじめに

インターネットは、さまざまな情報にアクセスするための基本的な手段であり、今や我々の生活に不可欠な情報ツールとなりつつある。インターネットとは、ごく簡単に言えば、世界規模のコンピューターネットワークのことである。Crystal（2001）によれば、インターネットは 1960 年代にアメリカで試験的ネットワーク（a network）として導入され、すぐに軍隊、連邦政府、大学、ビジネス分野等におけるユーザー、そして個人ユーザーへと拡大を見せた。このことからわかるように、INTERNET[1] という語は英語史のなかでもごく最近誕生した「新語」であるが、ここ数十年ですでに世界的な定着語となってきた。[2]

語の定着度をはかるにはいくつかの指標がある。社会的な認知度を示す指標としては、辞書への採録の有無や理解度・使用度の高さなどが挙げられる。手始めに Merriam-Webster.com Dictionary にて INTERNET を引いてみると、検索日（2022 年 8 月）の時点で見出し語は Internet（語頭が大文字表記）で採録されている。ただし、その他の表記変異形として internet（語頭が小文字表記）もあることが示されたうえで、以下のように定義される。

> Internet (noun)
> : an electronic communications network that connects computer networks and organizational computer facilities around the world — used with *the* except when being used attributively
>
> NOTE: In U.S. publications, the capitalized form *Internet* continues to be more common than *internet*, although the lowercase form is rapidly gaining more widespread use. In British publications, *internet* is now the more common form.
> — compare WORLD WIDE WEB
>
> （“internet”, Merriam-Webster.com Dictionary）

ここでは、INTERNET は複数のコンピューターネットワークや組織コンピューター設備をつなぐ世界規模の電子通信網と定義され、限定修飾的（attributive）用法を除き the を伴うとされる。注目したいのは、NOTE の部分である。それによれば、アメリカの出版物においては、大文字表記 Internet のほうが小文字表記 internet よりも一般的であること、それにもかかわらず小文字表記 internet の使用が拡大していること、またイギリスの出版物においては今日では小文字表記 internet のほうが一般的であることが述べられている。

ちなみにこの語義に続く What made you look up this word? というセクションでは、当該語を

検索したユーザーたちの書き込みを見ることができる。以下にそのうちのいくつか（sort by Best に指定した時の上位 3 位）を紹介する。

a . just to see if it is always capitalized. (21 May 2011)
b . To capitalize or not (7 September 2016)
c . I was confirming that you indeed capitalize it when using it as a proper noun and not a common noun. (1 April 2011)

（“internet”, Merriam-Webster.com Dictionary）

ここから見えてくるのは、INTERNET の語頭が大文字表記かそうでないか（以降、単に「INTERNET の大文字表記・小文字表記」と称することがある）の判断に迷う人が多いらしいということである。一般的には、太陽や月のようなこの世にただひとつだけの存在を指す場合、固有名詞としての用法が適用されるため、the Sun や the Moon のように大文字表記するべきとされる。INTERNET が一般に the Internet と大文字表記されてきたのは、同様の理由により固有名詞扱いされてきたからといえる。上記の書き込みのうち、最後のユーザー（c）による、「固有名詞の場合は大文字表記で、普通名詞の場合は小文字表記でよいのか」というコメントはこれと関連する。語彙の定着過程においては固有名詞の普通名詞化[3] が起こることがあり、それに伴い、例えば the Sun が the sun となるように形態的には脱大文字化が起こることがある。

このように、INTERNET の表記にはゆれがある。そこで本稿では、この語の表記の変異と変化について、コーパスを基盤とした変異研究の手法（Kuya 2021 等）を用いて実証的に論じる。まずは辞書やスタイルブックにおける表記の取り扱いについて概観したあと、コーパスから抽出した実時間データで表記の変異と変化を見ていく。

2．辞書・スタイルガイド等における表記の取り扱い

2.1. Crystal（2001）

Crystal（2001）は、著書 Language and the Internet の中で、一貫して大文字表記の（the）Internet を使用している。また、当該語の表記に関わる以下のような言及をしている。

… but there is no denying the unprecedented scale and significance of the Net, as a global medium. The extra significance is even reflected in the spelling, in languages which use capital letters: this is the first such technology to be conventionally identified with an initial capital. We do not give typographical enhancement to such developments as ‘Printing’, ‘Publishing’, ‘Broadcasting’, ‘Radio’, or ‘Television’, but we do write ‘Internet’ and ‘Net’.（下線は筆者による）

（Crystal 2001：3）

Crystal（2001）は、インターネットと同義のネット（the Net）のもつ、前例のない規模と「地球規模の情報媒体」（a global medium）としての重要性を強調した上で、このような特性が、この語

が慣習的に大文字で始まる（conventionally identified with an initial capital）ことと関連することを示唆する。つまり、インターネットやネットはその地球規模の情報媒体としての特性ゆえに慣習的に大文字表記される初めての技術であるということである。それに対して、印刷（printing）、出版（publishing）、放送（broadcasting）、ラジオ（radio）、テレビ（television）などの技術に対しては、少なくとも今日では、このような「印刷上の強調」（typographical enhancement）は施されない（つまり語頭が大文字表記されない）。

2. 2. Oxford English Dictionary Online（OED Online）（2000 年～）

OED Online[4] では、現在小文字表記 internet が見出し語として登録されているが、その他の表記変異形として大文字表記形も存在することが示されている（2022 年 8 月時点）。語源的には名詞の internetwork（*n.*）が縮まったものとして、以下のように定義されている。

> internet, n.
> 1. Originally (with lower-case initial): a computer network comprising or connecting a number of smaller networks, such as two or more local area networks connected by a shared communications protocol; an internetwork; *spec.* such a network (called *ARPANET*) operated by the United States Department of Defense. In later use (usually *the internet*): the global network comprising a loose confederation of interconnected networks using standardized communication protocols, which facilitates various information and communication systems such as the World Wide Web and email. Also: the resources accessible via this global network, *esp.* the World Wide Web. In early use usually *attributive*（下線は筆者による）
>
> （“internet”, OED Online）

上記では、もともとこの語は語頭が小文字で始まり複数の小規模ネットワークを相互に接続するコンピューターネットワークのひとつ（a computer network）を意味していた（つまり普通名詞扱いであった）が、のちに the をつけて the internet の形で the World Wide Web や email をはじめとするさまざまな情報・通信システムを促進する地球規模のネットワーク（the global network）を指すようになったこと、初期の用法においては一般に限定修飾的であったことが述べられている。

検索日時点で OED Online に掲載されている用例を見てみると、初出年は 1974 年となっている[5]。掲載された用例のうち最も古いものと最も新しいもの 2 例ずつを以下に掲げる。1974 年の他の名詞を修飾する限定修飾的用法においては小文字表記（internet transmission control program）であるが、1980 年、2004 年、2013 年の the を伴う用法においては小文字表記と大文字表記の両方の用例が存在することがわかる。

> 1974 V. Cerf et al. *Request for Comments* (Network Working Group) (Electronic text) No. 675. 1 (*title*) Specification of internet transmission control program.
> 1980 J. B. Postel in *IEEE Trans. Communications* Apr. 607/2 The IP is a datagram protocol. The collection of interconnected networks is called an internet. IP is the

network protocol of the internet... To provide that type of end-to-end reliable ordered delivery of data the ARPA internet uses TCP

2004 *Webactive* 14 Oct. 67 When you connect to the internet or if you use a network, all traffic will be monitored and checked for suspicious and undesirable programs.

2013 *N.Y. Rev. Bks.* 7 Nov. 82/4 Anything the police do might be on the Internet within minutes and visible to uncountable numbers of virtual bystanders.（下線は筆者による）

（"internet", OED Online）

ちなみに、OED に見出し語 INTERNET が追加されたのは、OED がオンライン版である OED Online として公開された年の翌年 2001 年 6 月のことである（Dent 2015）。その技術の急速な普及・発展を背景に Internet（*n.*）と internet（*v.*）の新語としての追加が急がれたことがうかがえる。特筆すべきは、OED Online に見出し語として追加された 2001 年の時点では動詞が小文字表記 internet（*v.*）であるのに対し、名詞は大文字表記 Internet（*n.*）となっていることである（OED Online n.d., New words list June 2001 を参照）。つまり、OED Online では採録当初は大文字表記が採用されていたが、のちになって小文字表記に改訂されたということである。表記の改訂がどの時点で起きたかについて厳密にはトレースできないが、OED 補助編集者である Dent（2015）の記事では大文字表記 the Internet が一貫して使われていることから改訂はそれ以降であると推察される。

2. 3. スタイルガイド（AP Stylebook、Chicago、APA、MLA）

2. 3. 1. Associated Press Stylebook（AP Stylebook）

報道関係者や一般向けの AP Stylebook は、2016 年版から、小文字表記の internet と web を採用することを発表した（Easton 2016）。以下に示す通り、2016 年版では見出し語に小文字表記が採用され、巻頭では最近の傾向と AP Stylebook の公式辞書の方針転換に合わせて小文字表記を採用する旨が述べられている。AP Stylebook は 1985 年版から 2020–2022 年版まで毎年改定を行っており、後述する学術系のスタイルガイドよりも改訂頻度が高い。よって AP の方針変更は他の学術系スタイルガイドにも一定の影響を与えた可能性がある。

One of the main changes is *internet*. We now spell it lowercase, reflecting a growing trend and a change by our official dictionary, Webster's New World College. We have also made *web* lowercase in all instances, and *webpage* and *webfeed* one word.

（Associated Press 2016：v）

2. 3. 2. The Chicago Manual of Style

これに次いで The Chicago Manual of Style は、2017 年刊行の第 17 版から、小文字表記 internet を採用すること、e-mail を email（ハイフンなし）にする方針を打ち出した（Chicago Manual 2017）。2010 年刊行の第 16 版では巻末インデックスにおける見出し語は大文字表記 Internet となっているが（University of Chicago 2010：971）、第 17 版では小文字表記 internet となっている（University of Chicago 2017: 1070）。

2.3.3. Publication Manual of the American Psychological Association（APA）

APA は、2010 年刊行の第 6 版で巻末インデックスの見出し語に INTERNET を採録しているが、ここでは全ての見出し語の語頭が大文字表記されているためマニュアル本文中を探すと、大文字表記 the Internet が採用されていることがわかる（APA 2010：236）。しかし、2019 年 10 月 3 日、Twitter 上で小文字表記 the internet と the web、そして email（ハイフンなし表記）の使用を推奨する旨を発表しており（APA 2019）、2020 年刊行の第 7 版では見出し語から INTERNET が削除されている。ちなみに、関連語 webpage や website は本文において小文字表記である（APA 2020：298）。

2.3.4. Modern Language Association Handbook（MLA Handbook）

MLA Handbook は、2016 年刊行の第 8 版においてインデックスの見出し語に大文字表記 Internet を採用しており（MLA 2016：137）、2020 年の時点でも大文字表記を推奨している（MLA Style Center 2020）が、その後の方針については最新版で確認するよう促している。そこで 2021 年刊行の第 9 版（MLA 2021）を確認すると、APA と同様、見出し語そのものが削除されていることから、少なくとも INTERNET の大文字表記について積極的ではなくなったことがうかがえる。見出し語として採録されている関連語 webpage や website には小文字表記が採用されていることから（MLA 2021：366-367）、他のスタイルガイドと同様 INTERNET についても小文字表記に移行していくものと推測される。

3．リサーチクエスチョンと研究方法

3.1. 本稿の問いと予測

辞書記述は、英米英語いずれにおいても、INTERNET の表記にはゆれのあることを示している。また、上述したスタイルガイドの動きは、INTERNET（とその他のインターネット関連用語）の表記の脱大文字化を促進するものであるため、大文字表記から小文字表記への交替が進行していることが予測される。そこで本稿では、(1) 脱大文字化が進行中であるか、(2) スタイルガイドの方針変更時期と変化速度にどのような関連があるか、(3) この変化にレジスター差は見られるか、という 3 つの問いを立てる。変化速度については、辞書やスタイルガイドの方針転換の動きを踏まえると、全体として 2010 年代後半以降に大文字表記から小文字表記への交替が加速するとの予測がたつ。

3.2. コーパスの選定

本稿ではコーパスを基盤とした変異研究（Kuya 2021 等，Szmrecsanyi 2017 も参照）の手法を用いて、実時間データから上記の 3 つの問いについて検証する。テキストコーパスは、語彙レベルの変異（久屋 2013；Kuya 2019）や形態に関わる変異（南部 2007；佐野 2008）に加え、今回のような表記レベルの変異の観察に向いている。今回利用するのは、2022 年 8 月時点で 19 の英語オンラインコーパスへのインターフェイスを提供する English-Corpora.org（https://www.english-corpora.org/）である。数百万語規模から百億語規模まで、複数の英語変種、時代、ジャンル（genre）にまたがるさまざまな大規模コーパスが存在するため、目的に応じてコーパス選定を行う必要がある。

代表的な通時的コーパス（diachronic corpus）としては The Corpus of Historical American

English（COHA）（Davies 2010）が挙げられる。このコーパスは過去 200 年間（1820 ～ 2019 年）のアメリカ英語約 4.8 億語からなる。（1）TV/Movies、（2）Fiction、（3）Popular Magazines、（4）Newspapers、（5）Non-fiction Books の 5 つのジャンル[6] を含む均衡コーパス（balanced corpus）とされているが、すべての年代を通じて Fiction が半分近く（40 ～ 50%）を占めていること、また TV/Movies と Newspapers のデータがそれぞれ 1930 年代以降と 1860 年以降のものに限られることから、ジャンル間の比較に使用する際には注意が必要である。今回は INTERNET の登場時期と新変異形 internet の使用割合の通時的変遷を概観するために利用する。なおインターネット技術（の原型）が登場したのが 1960 年代ごろである（Crystal 2001）とすると、1960 年代からスタイルガイドの方針転換があった 2010 年代までを網羅するコーパスが必要であるが、COHA はこの条件によく適合する。この他に TV Corpus や Movie Corpus もこの条件を満たすコーパスではあるが、書きことばのジャンルをより幅広く観察できる COHA を選択する。ちなみにこれらは TV/Movies というひとつのサブジャンルとして COHA に含まれている。

代表的な共時的コーパス（synchronic corpus）としては The Corpus of Contemporary American English（COCA）（Davies 2008-）が挙げられる。このコーパスは過去 30 年間（1990 ～ 2019 年）のアメリカ英語約 10 億語からなり、（1）Blogs、（2）Web Pages、（3）TV/Movies、（4）Spoken、（5）Fiction、（6）Popular Magazines、（7）Newspapers、（8）Academic Journals の 8 つのジャンル[7] からバランスよく抽出された均衡コーパスである。COHA に比べてジャンル数が豊富で、各ジャンルの合計収録語数が 1.2 ～ 1.3 億語に調整されているため、ジャンル間の比較に適している。また、各年の収録語数も約 2,500 万語ずつに調整されているため、ジャンルごとの経年比較も可能で、INTERNET のようにここ数十年の間に登場した新語の研究においては通時コーパスとしての利用価値も高いと考えられる。今回は新変異形 internet の使用割合の通時変化をジャンルごとに観察するために利用する。ただし、Blogs と Web Pages は 2012 年 10 月に取得されたテキストを収録したもので、検索結果には年代情報が 2012 と表示されるものの実際には 2012 年よりも前にウェブ上に掲載されたテキストも含まれるため、通時的分析に利用する際には注意が必要である。

3.3. 検索式の検討

今回は固有名詞として the に後続する INTERNET の表記の変異を調べるため、大文字表記 the Internet と小文字表記 the internet を抽出する。

第 1 節から、INTERNET については固有名詞扱いの場合大文字表記、普通名詞扱いの場合小文字表記という、品詞と表記に対応関係の存在することがわかる。ここで問題となってくるのが、COHA や COCA において INTERNET の品詞タグづけがどのように行われているかということである。そこで、COHA にて、①特に品詞指定のない文字列による語検索（the Internet, the internet）と、①に品詞条件を加えた 3 種類の品詞検索：②固有名詞（the internet.[np*]）、③普通名詞（the internet.[nn*]）、④名詞一般（the internet.[n*]）[8] の合計 4 パターンの検索を行った。検索モードは List を指定し、CASE SENSITIVE のオプションを YES にすることで大文字表記（the Internet）と小文字表記（the internet）のヒット数を別々に集計する方法を用いた。

検索結果を表 1 に示す。①品詞指定がない the Internet と the internet を検索すると、ヒット数は合計 3,048 例であった。品詞指定をした場合、②固有名詞としてタグづけされている例はわずか

（133 例）で、③普通名詞としてタグづけされているケースがほとんど（2,904 例）であった。しかし、普通名詞に限定した検索結果には大文字表記と小文字表記が混在しており、これは「普通名詞としての用法は小文字表記である」という対応関係と矛盾する。さらに、普通名詞としてタグづけされた小文字表記（the internet）の用例のなかに地球規模のネットワークを意味する固有名詞用法がいくつも発見された。よって、第 1 節で述べたような品詞と表記の対応関係が本コーパスにはないことから、今回は固有名詞・普通名詞の区別をしない、④名詞一般を指定する検索式（the internet.[n*]）[9] を利用することとした。なお、④名詞一般のヒット数（3,037 例）は、②と③のヒット数の合計と一致する。

表 1 ：検索式の検討に使用した検索フォーマットとヒット数（COHA）

検索種	検索フォーマット	Internet	internet	TOTAL
①文字列（品詞指定なし）	the Internet と the internet	2,677	371	3,048
②固有名詞に指定した場合	the internet.[np*]	133	0	133
③普通名詞に指定した場合	the internet.[nn*]	2,538	366	2,904
④名詞一般に指定した場合	the internet.[n*]	2,671	366	3,037

4. 結　　果

4.1. INTERNET の登場時期と新旧変異形の分布の通時的変遷（COHA）

まず、INTENRNET の登場時期を見るために、3.3 節に示した検索（検索式 the internet.[n*]、検索モードは List、CASE SENSITIVE は YES に設定）を COHA で行った。その結果、合計 3,037 例（大文字表記 2,671 例、小文字表記 366 例）がヒットした（表 2）。結果は 10 年区切りで表示されているが、INTERNET の使用が全くない年代は除いた。

まず、1980 年代までは用例数自体が 1 桁と少ない[10] ものの、1990 年代以降に急増し、その後 2010 年代に至るまで 1,000 例前後の頻度を維持している。このことから、INTERNET の使用が一般的になったのは、1990 年代であるといえる。次に新旧変異形の通時的変遷を見てみると、1990 年代以降の小文字表記率（「internet 率」[11] とする）が、1990 年代から 2000 年代までは 2.4% ～ 4.0% と非常に低い水準にとどまるが、2010 年代に一気に 30.8% にまで急増する。このことから、スタ

表 2 ：COHA における the Internet / the internet 使用頻度

変異形	1890s	1930s	1940s	1970s	1980s	1990s	2000s	2010s	TOTAL
Internet	1		1		1	712	1,256	700	2,671
internet		1		2		16	51	296	366
TOTAL	1	1	1	2	1	728	1,307	996	3,037
internet 率	0%	100%	0%	100%	0%	2.4%	4.0%	**30.8%**	12.6%

イルガイドの方針変更の時期にあたる 2010 年代に the Internet から the internet への交代が加速したことが確認できる。

4.2. ジャンル別にみる新旧変異形の分布と通時的変遷（COCA）

続いて、COCA にも同じ検索（検索式 the internet.[n*]、検索モードは List、CASE SENSITIVE は YES に設定）を行ったところ、表 3・表 4 に示す結果となった。the INTERNET や the INternet など、語頭以外も大文字になっている用例は除いた。

表 3 は、COCA における the Internet / the internet の分布をジャンル別に示したものである。8 つのジャンル区分のうち、internet 率（%）は Blogs［BLOG］で 58.8% と最も高く、Web Pages［WEB］の 38.9%、TV/Movies［TV/M］の 30.9% と続く。残りのジャンルは全ジャンル合計の平均値（28.2%）より低く、最も低いのが Newspapers［NEWS］の 3.8%、続く Popular Magazines［MAG］、Academic Journals［ACAD］、Spoken［SPOK］はいずれも十数パーセントである。つまり internet 率は、話しことばの書き起こしである TV/Movies や Spoken を除くとインターネット媒体（Blogs や Web Pages）で高く、報道関連分野（Newspapers や Popular Magazines）と学術分野（Academic Journals）においては比較的低いことが指摘できる。Blogs と Web Pages についてはコーパスの特性上書かれた年が厳密には特定できないものの、これらのジャンルのテキスト収集が行われたのは 2012 年である。したがって少なくとも 2012 年の時点、つまり主要スタイルガイドの方針改訂が起こる前の時点で、すでに他のジャンルよりも変化がかなり進行した状態であったことは読み取れる。

表 4 は、COCA における the Internet / the internet の分布の通時的変遷を示したものである。なお、この表には年代情報が特定できない Blogs と Web Pages からの用例は含まれていない（よって表 4 の用例数の合計は 49,222 − 10,480 − 10,526 = 28,216 となる）。COHA で確認した通り、INTERNET は 1990 年代以降に集中して出現しているため、1990 年から現在までのわずか 30 年間のコーパスであ

表 3：COCA における the Internet / the internet の共時的分布（ジャンル別）

変異形	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	TOTAL
Internet	**4,321**	**6,432**	1,353	5,382	929	6,382	6,066	4,483	35,348
internet	**6,159**	**4,094**	606	896	265	934	238	682	13,874
TOTAL	**10,480**	**10,526**	1,959	6,278	1,194	7,316	6,304	5,165	49,222
internet 率	**58.8%**	**38.9%**	30.9%	14.3%	22.2%	12.8%	3.8%	13.2%	28.2%

表 4：COCA における the Internet / the internet の通時的変遷（BLOG と WEB は除く）

変異形	1990-4	1995-9	2000-4	2005-9	2010-4	2015-9	TOTAL
Internet	406	5,926	6,700	4,943	3,942	2,678	24,595
internet	17	96	201	694	671	1,942	3,621
TOTAL	423	6,022	6,901	5,637	4,613	4,620	28,216
internet 率	4.0%	1.6%	2.9%	12.3%	14.5%	**42.0%**	12.8%

るCOCAであっても、当該語の通時的変遷を観察するのに有用である。COCAでは、COHAよりも短い5年区切りのスパンで通時的変遷を概観することができるが、それによればINTERNETの出現回数は1990年代前半までは500例未満と僅かだが、1990年代後半には一気に6,000例に到達し、2000年代前半には7,000例にせまりピークを迎える。それ以降は用例数としては若干落ち込んでいるが、5,000例前後を維持している。internet率でみると、2000年代前半まではわずか数％にとどまるが、2000年代後半以降は10%を超え、2010年代後半に一気に40%を超える。

この結果は上述のCOHAの調査結果と矛盾しない。ここでもthe Internetからthe internetへの交替が進んでいることが確認できた。さらにCOHAよりも短いスパンでみることにより、新変異形の出現率が急上昇したのはスタイルガイドの方針変更の時期にあたる2010年代後半であることが特定できた。

さらに表4が示す通時的変遷のジャンル差を見るためにジャンル別に検索（検索式the internet.［n*］、検索モードはList、CASE SENSITIVEはYESに設定）・集計したところ、ジャンルにより異なる変化の動向を示した（図1）。まず、internet率は、TV/Movies［TV/M］やFiction［FIC］では1990年代の終わりまでに一定の高さに達していることが確認できる。Spoken［SPOK］では1990年代前半の乱れ（用例数が他の年代よりも少なかったことことに起因する可能性がある）を除くと、2000年代後半から徐々に採用率を伸ばしている。Academic Journals［ACAD］では、2000年代後半の一時的な増加傾向をのぞけば、2010年代前半以降緩やかに新変異形への交替が進んでいるとみることができる。

注目したいのは2010年代前半から後半の間に新変異形への交替が急激に進んだPopular Magazines［MAG］とNewspapers［NEWS］である。Popular Magazinesは2010年前半時点で新変異形の採用率が6.6%で、最も保守的な情報媒体のうちのひとつであったのが、2010年後半になるとそれが50％を超え1位に躍り出ている。Newspapersも2010年代前半と後半を境に新変異形の採用率が1.1%から40％近くにまで急増した。それ以外のジャンルではスタイルガイドの方針が転換される2010年代前半よりも前からある程度変化が進行しはじめており、方針転換後に変化が大きく加速したわけでもない。例えば、前述の通りTV/MoviesやFictionは早い段階（1990年代）から変化が進行していたジャンルのひとつであるが、スタイルガイドの方針転換後の変化速度（グラフの傾き）にさほど変わりがない。

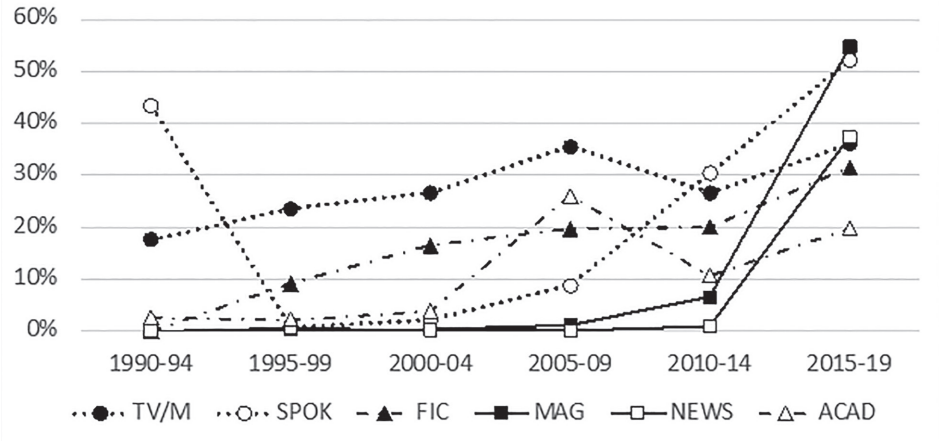

図1：COCAにおけるinternet率（%）の通時的変遷（ジャンル別）

5．おわりに：考察と今後の課題

本稿では、言語変異・変化研究の観点から、新語 INTERNET の 2 つの表記変異形のコーパスにおける分布を通時的・共時的側面から概観した。調査の結果、(1) the Internet から the internet への交替があらゆるジャンルで進行中であること、(2) コーパス上で小文字表記の出現率が急増するのは 2010 年代（後半）であり、それはスタイルガイドで当該語の表記に関する方針が転換されはじめた時期に対応していること、(3) 新変異形への移行の開始時期や全体的な採用率はジャンルによって異なることが明らかになった。最後の点は、個々のジャンルの特徴について考察するうえで有益な示唆を含んでいる。

まず、Popular Magazines や Newspapers では、AP Stylebook 等主要スタイルガイドの方針転換期（2010 年代後半）に新変異形の採用率がほぼ 0 % から 40 ～ 50% にまで急増した点が特徴的であった。考えられる可能性としては、これらの媒体がスタイルガイドの影響を最も大きく受けたか、もしくはこれらの媒体がスタイルガイドの方針転換に影響を及ぼしたのかもしれないということである。スタイルガイドの方針が転換されるまでは TV/Movies や Fiction や Spoken が変化をリードしていたものの、今後は Popular Magazines や Newspapers が変化をリードしていくであろうことが予測できる。

次に、Blogs や Web Pages で新変異形の出現率が著しく高かったことから、これらのジャンルの特性について考える必要がある。Crystal（2001）は Netspeak と呼ばれるインターネット上のことばづかいにおいてキャピタライゼーションの原則は一般に守られないと指摘する。さらに Netspeak において語頭の大文字表記が避けられることに「キーストロークを減らせ」(Save a keystroke）というスタイル上の掟が関わっているという [12]。Blogs や Web Pages における表記の変化はこのような制約を背景に進行した可能性がある。これら 2 つのジャンルで INTERNET の用例数がそれぞれ 1 万例以上と最も多かったこと（表 3）も併せて考えると、インターネット上のプラットフォームが今回の変化をリードし、結果としてスタイルブックの方針変更に影響を与えた可能性も否定できない。

このような機器操作上の理由に加え考えなければならないより本質的な問題は、インターネットやウェブという媒体がことばの実践にどのような影響を与えうるかということである。Crystal（2001：Ch.2）は Netspeak が話しことばか書きことばか（speech or writing）という問いを立て、それが話しことば的でもあり書き言ことば的でもあることを示した上で、単純な足し合わせではない、コミュニケーションの新種（a new species of communication）としてみるべきであると指摘する。インターネット上のことばをレジスターやスタイルといった既存の共時的変異の分析にどう位置付けるかは今後の課題であり、それが言語変異・変化研究にどのような知見をもたらすかは、今後の研究の蓄積を待たなければならないだろう。

その他の課題として、話しことばの書き起こしである TV/Movies や Spoken のジャンル特性の分析が挙げられる。これらのジャンルにおける字幕表記や書き起こしに関わるルールなどの観点から新変異形の出現率との関連についてさらに詳しく考察する必要があるだろう。また、Academic Journals で 2000 年代後半に internet 率が一時的に増加した理由の解明と、本稿で十分に検証できなかった、今日では小文字表記が一般的であるというイギリス英語の変異・変化についても今後追

求していく必要がある。さらに、スタイルガイドの方針変更という言語外的要因に加え、言語内的要因についても検討する必要がある。前述した 2019 年 10 月 3 日の APA の方針変更に対して、翌日に Twitter 上で以下のような批判的な反応があった。

> Sorry I think the small i is stupid. There is more than one internet. Internet refers to the basic www. version that most know and interact with. But, for example, there is www3. which is science based, the "dark net" etc. Internet - capital I has a meaning why do you mess w/it?
>
> （[@DrDebD]2019-10-4, Twitter）

つまり、internet に当たるものはいくつかあるが、Internet とすれば最も認知度・使用度の高い www.（the World Wide Web）のことであるとわかるため、大文字表記のままでよいのではないか、その他の internet、例えば www3. やダークネット（dark net）と呼ばれるネットワークとの区別を可能にするものが大文字 I である、という意見である。the Internet と the World Wide Web は厳密には全く同じものを指すわけではない（Associated Press 2016：300）が、一般的には両者は同じ意味で使われることも多く、そのような事情がこうした反応の背景にあると思われる。

the Internet から the internet への表記の変化はターニングポイントを迎えたが、いまだ過渡期にある。スタイルガイドの方針変更（言語外的要因）や意味の差異化の視点（言語内的要因）がこの言語変化にどのような影響を与えるか、言語変異・変化研究の観点からも興味深い。さらに、変異研究では比較的新しいレジスターであるインターネットやウェブという存在そのものが言語変化一般に与える影響も、今後検討されていくべきであろう。

注

1）本稿では、表記レベルの変異（variation）を扱う関係上、曖昧性を避けるため以下の 3 つの表記を使い分ける。個々の表層形、すなわち変異形（variant）を指す場合は the Internet または the internet と表記する（文脈によっては the をつけないで表記することもある）。これら 2 つの変異形を内在する語の基本形、すなわち言語変数（variable）を指す場合は INTERNET と表記する。つまり、言語変数としての INTERNET は the Internet または the internet のいずれかの表記変異形として出現する。

2）非英語圏では借用語として移入されている。音訳借用の例としてはドイツ語の (das) Internet、フランス語の (l') Internet、日本語の「インターネット」、韓国語の인터넷（in-teo-nes）、翻訳借用の例としては中国語の「互联网」（hù-lián-wǎng）などがある。

3）例えば、Escalator や Videotape などの商標名の普通名詞化や、(the Forth Earl of) Sandwich（第 4 代サンドウィッチ伯爵）などの人名の普通名詞化（Östberg 1905）も含まれる。このとき語頭が脱大文字化（escalator、videotape、sandwich）することがあるが、Solomon（聖書上の人名に因んだ賢者の意）や Frankenstein（Mary Shelley の作品中の登場人物名に因んだ邪悪の化身の意）など、大文字表記のまま普通名詞的に使われる例もある。脱大文字化はインターネット関連用語であるウェブ（Web ＞ web）などにも起こっている。その他、インターネット関連用語の定着過程においては、接辞化（electronic ＞ e-）、ハイフネーションの消失（e-mail ＞ email）なども起こっている。（2. 3 節参照）

4）OED は、語のつづり・語形・発音・意味・用法の変異や変化、語の初出年などを、実際の用例に基づいて記述した辞書で、語の共時的変異と通時的変化の観察が可能である。OED のオンライン版である OED Online は、OED 第 2 版（1989 年刊行）と Additions Series とよばれる第 2 版の増補全 3 巻（Volume 1–2 は 1993 年、Volume 3 は 1997 年刊行）をベースとして 2000 年 3 月にオンライン公開が開始された。それ以来、毎年新語の追加が行われており、新しく追加された語は Updates to the OED（https://public.oed.com/updates/）から

確認できる（随時更新される）。

5）初出年の認定は辞書によって異なり、更新されることがある。インターネットのルーツがアメリカで 1960 年代に登場した試験的なネットワークである（Crystal 2001）ことを考慮すると、この語自体が使用され始めたのは 1974 年よりも早い時期である可能性もある。Merriam-Webster.com Dictionary は、現在の形である地球規模のネットワークの意味での初出年を 1986 年としている。

6）（1）TV/Movies はテレビ・映画の字幕をテキスト化したもの、（2）Fiction はフィクション書籍（映画や演劇のスクリプトを含む）、（3）Popular Magazines は一般雑誌、（4）Newspapers は新聞、（5）Non-fiction Books はノンフィクション書籍である。TV/Movies は 2021 年の更新で COHA の 4 つの既存ジャンルに新たに追加されたジャンルである。ちなみにこの更新では他の年代に比べ規模が小さくジャンル間のデータ量のばらつきが大きい 1810 ～ 1819 年のデータが削除され、2010 ～ 2019 年のデータが追加補充された。

7）（1）Blogs は、英語使用国 20 カ国のウェブテキストから約 19 億語を収録した Global Web-Based English（GloWbE）と呼ばれるコーパスのアメリカ英語部門の Blog ジャンルを構成するサブセットで、（2）Web Pages は同じく GloWbE アメリカ英語部門の General ジャンルを構成するサブセットである。この 2 つのジャンルを利用する際の留意点としては、GloWbE の収録テキストは 2012 年 10 月に取得されたウェブテキストである（つまり厳密には年代情報が特定できない）こと、また Web Pages には Blogs からのテキストが含まれている（一部重複がある）可能性のあることである。その他、（4）Spoken はテレビ・ラジオ番組の自然発話の書き起こし、（8）Academic Journals は学術論文である。（3）TV/Movies、（5）Fiction、（6）Popular Magazines、（7）Newspapers の一部は COHA の内容と重複する。

8）長谷部（2020）を参考に、品詞検索では CLAWS7 タグセットを用いた。アステリスク（*）は 0 文字以上の文字を表すワイルドカードで、[n*] は [np*] と [nn*] を含む。

9）ちなみに COHA の場合検索式 the internet.[n*] は INTERNET が名詞を修飾する限定修飾的用法で使われるケースも抽出する。そのようなケース（the internet.[n*] [n*] で検索するとヒット数は 231 例）を見てみると、限定修飾的用法にもかかわらず大文字表記の使用例が多く、Internet age と internet age のように大文字表記と小文字表記が混在するケースも少なくない。固有名詞（the Internet Archive や the Internet Explore）が一部含まれる点には留意が必要であるものの、全体的には限定修飾的用法にも表記レベルの変異が存在するため、今回の分析対象から除外しないこととした。

10）インターネット技術（の原型）が誕生したのが 1960 年代以前である（Crystal 2001）とすると、それ以前の用例は注意が必要である。例えば 1890 年代（Fiction）の用例は引用元情報にあたる The Internet Archive/American Libraries という語句が検索にかかった結果であるので、除外して考えるべきである。

11）「internet 率」は、the internet／（the Internet＋the internet）×100 で算出した。例えば 1990 年代の internet 率は 16／（712＋16）×100＝2.4% となる。

12）さらに「疑わしきはツメよ」（When in doubt, close it up）という掟のもとでは、分かち書きされていた複合語がハイフネーションされ、さらに次の段階でハイフンが取れることも指摘している（e-mail ＞ email）。

参考文献

APA (2019). Preferred spellings. APA. https://apastyle.apa.org/style-grammar-guidelines/spelling-hyphenation/preferred-spellings（参照 2022-8-23）.

Chicago Manual (2017-3-28) Announcing *the Chicago Manual of Style*, 17th Edition. Chicago Manual. https://cmosshoptalk.com/2017/03/28/announcing-the-chicago-manual-of-style-17th-edition/（参照 2022-8-22）.

Crystal, David (2001). *Language and the Internet*. Cambridge University Press.

Dent, Jonathan (2015-6-25). Release notes: capturing the interweb of words. Oxford University Press. https://public.oed.com/blog/june-2015-update-release-notes-capturing-the-interweb-of-words/（参照 2022-8-22）.

DrDebD Furmom to Buddy, Lulu, Belle, Merlot #ZSHQ [@DrDebD] (2019-10-4). *Sorry I think the small i is stupid. There is more than one internet. Internet refers to the basic www*. [Tweet]. Twitter. https://twitter.com/DrDebD/status/1179848202520018944（参照 2022-8-23）.

Easton, Lauren (2016-4-2). Ready to lowercase 'internet' and 'web'. Associated Press. https://www.apstylebook.com/blog_posts/5（参照 2022-8-22）.

Kuya, Aimi (2019). *The Diffusion of Western Loanwords in Contemporary Japanese: A Variationist Approach*. Hituzi Syobo Publishing.

Kuya, Aimi (2021). A corpus-based variationist approach to the use of *it is I* and *it is me*: A real-time observation of a syntactic change nearing completion in COHA. *Gengo Kenkyu*, 159: 7–35.

MLA Style Center (2020-9-25). Is *Internet* capitalized in MLA style?. *Modern Language Association of America*.

https://style.mla.org/capitalization-of-internet/（参照 2022-8-23）.
OED Online (n.d.). New words list June 2001. Oxford University Press. https://public.oed.com/updates/new-words-list-june-2001/（参照 2022-8-22）.
Östberg, H. O. (1905). *Personal Names in Appellative Use in English*. Almqvist & Wiksell. 鈴木重成（訳）（1958）『人名の普通名詞化』研究社 .
Szmrecsanyi, Benedikt (2017). Variationist sociolinguistics and corpus-based variationist linguistics: Overlap and cross-pollination potential. *Canadian Journal of Linguistics / Revue canadienne de linguistique*, 62 (4): 685-701.
久屋愛実（2013）「現代書き言葉における外来語の共時的分布 :「ケース」を事例として」『国立国語研究所論集』6：45-65.
佐野真一郎（2008）「『日本語話し言葉コーパス』に現れる「さ入れ言葉」に関する数量的分析」『言語研究』133：77-106.
南部智史（2007）「定量的分析に基づく「が／の」交替再考」『言語研究』131：115-149.
長谷部陽一郎（2020）「English-Corpora.org を用いた言語データの採取」（Ver. 2020-05-29）. https://yohasebe.com/assets/docs/ECO-Tutorial.pdf（参照 2022-7-29）.

辞書

"Internet". *Merriam-Webster.com Dictionary*. Merriam-Webster. https://www.merriam-webster.com/dictionary/Internet（参照 2022-8-22）.
"internet, n.". *OED Online* (2022-6). Oxford University Press. https://www.oed.com/view/Entry/248411?rskey=b1Qzco&result=1（参照 2022-8-22）.

スタイルガイド

APA (2010). *Publication Manual of the American Psychological Association* (6th Edition). The American Psychological Association.
APA (2020). *Publication Manual of the American Psychological Association* (7th Edition). The American Psychological Association.
Associated Press (2016). *Associated Press Stylebook and Briefing on Media Law* (51st Edition, 2016). The Associated Press.
MLA (2016). *MLA Handbook* (8th edition). The Modern Language Association of America.
MLA (2021). *MLA Handbook* (9th edition). The Modern Language Association of America.
University of Chicago (2010). *The Chicago Manual of Style* (16th Edition). The University of Chicago Press.
University of Chicago (2017). *The Chicago Manual of Style* (17th Edition). The University of Chicago Press.

コーパス

Davies, Mark (2008-). *The Corpus of Contemporary American English (COCA)*. https://www.english-corpora.org/coca/（参照 2022-8）.
Davies, Mark (2010). *The Corpus of Historical American English (COHA)*. https://www.english-corpora.org/coha/（参照 2022-8）.

（本学文学部准教授）

Capitalization of INTERNET:
Is it *the Internet* (Capitalized) or *the internet* (Lowercased)?

by

Aimi Kuya

This paper discusses change in the capitalization of the "new word" INTERNET within the framework of corpus-based variationist linguistics. As reflected in its capitalized form (*the Internet*), this word has been conventionally treated as a proper noun; however, several major stylebooks have recently started to suggest that it should be lowercased (*the internet*). Such a modification of capitalization rules indicates a shift in the state of the word from a proper noun towards a common noun, possibly as a result of increased use of web-based technology. Observations of two corpora, COHA and COCA, provide diachronic and synchronic distributions of these variant forms: (i) the diffusion of the new, lowercased variant is in progress across genres; (ii) the proportion of the lowercased variant skyrocketed in the second half of the 2010s, which corresponds with the time when stylebooks began to change their policy on the capitalization of INTERNET; and (iii) the stylebooks' policy change finds a high rate of correlation in genres such as magazines and newspapers. In addition, the relatively high proportion of the use of the new variant in blogs and web pages raises a question as to how these web-based platforms should be treated in the analysis of synchronic variation according to register and style.