

あらうと想像される。今では遠い停車場の機關車の出し入れの音が時として非常に間近く聞えると云つたやうな現象と姿を變へて注意されるやうになつた。狸も大分モダン化したのである。此のやうな現象でも精細な記録を作つて研究すれば氣象學上に有益な貢獻をする事も可能であらう。

「天狗」や「河童」の類となると物理學や氣象學の範圍からは大分遠ざかるやうである。併し「天狗様の御囃子」などいふものは矢張前記の音響異常傳播の一例であるかも知れない。天狗和尚とジュースの神の鷲との親族關係は前に述べたが、河童が海龜の親類である事は善庵隨筆に載つて居る「寫生圖」と記事、又筠庭雜錄にある繪や記載を見ても明である。河童の寫生圖は明に龜の主要な特徴を具備して居り、其記載には現に「龜の如く」といふ文句が四箇所もある。さうだとすると此等の河童捕獲の記事は或る年の或る月に或る沿岸で海龜が獲れた記録になり、場合によつては海洋學上の貴重な參考資料にならないとは限らない。

序ながら印度邊の國語で海龜を「カチファ」といふ。「カツパ」と似て居て面白い。

尤も「河童」と稱するものは、其實色々雑多な現象の綜合されたものであるらしいから、今

日此れを論ずる場合にはどうしても一旦此れを其の主要成分に分析して各成分を一々吟味した後に、此等が如何に組み合はされて居るか、又時代により地方により其の結合形式が如何に變化して居るかを考究しなければならぬ。此れは中々容易でないが、もし出來たら可也に面白く有益であらうと思ふ。此のやうな分析によつて若干の化物の元素を析出すれば、他の化物は此等の化物元素の異なる化合物として説明されないと限らない。CとHとOだけの組合せで多數の有機物が出るやうなものかも知れない。此れも一つの空想である。

要するにあらゆる化物を如何なる程度迄科學で説明しても化物は決して退散も消滅もしない。唯化物の顔貌が段々にちがつたものとなつて現はれるだけである。人間が進化するにつれて、化物も進化しない譯には行かない。併しいくら進化しても化物は矢張化物である。現在の世界中の科學者等は毎日各自の研究室に閉籠り懸命に此等の化物と相撲を取り其正體を見破らうとして努力して居る。併し自然科學界の化物の數には限りがなく各々の化物の面相にも際限がない。正體と見たは枯柳であつて見たり、枯柳と思つたのが化物であつたりするのである。此の化物と科學者の戦は恐らく永遠に續くであらう。さうしてさうする事によつて人間と化物とは

永遠の進化の道程をたどつて行くものと思はれる。

化物がないと思ふのは却つて本當の迷信である。宇宙は永久に怪異に充ちて居る。あらゆる科學の書物は百鬼夜行繪卷物である。それを繙いて其の怪異に戰慄する心持がなくなれば、もう科學は死んでしまふのである。

私は時々密に想ふ事がある、今の世に最も多く神秘の世界に出入するものは世間からは物質科學者と呼ばれる、科學研究者ではあるまいか。神秘的あらゆるものは宗教の領域を去つて何時の間にか科學の國に移つてしまつたのではあるまいか。

又こんな事を考へる、科學教育は矢張昔の化物教育の如くすべきものではないか。法律の條文を暗記させるやうに教へ込むべきものではなくて、自然の不思議への憧憬を吹き込む事が第一義ではあるまいか。これには教育者自身が常に此の不思議を體驗して居る事が必要である。

既得の知識を繰返して受賣りするだけでは不十分である。宗教的體驗の少い宗教家の説教で聽衆の中の宗教家を呼びさます事は稀であると同じやうなものであるまいか。

斯んな事を考へるのは或は自分の子供の時に受けた「化物教育」の藥が利き過ぎて、折角受

けたオーソドックスの科學教育を自分の「お化鏡」の曲面に映して見て居る爲かも知れない。さうだとすれば此の一篇は一つの懺悔録のやうなものであるかも知れない。此れは讀者の判斷に任せる外にない。

傳聞するところによると現代物理学の第一人者である丁抹のニエルス・ボーアは現代物理学の根本に横はる或る矛盾を論じた際に、此の矛盾を解き得る迄に吾々人間の頭は未だ進んで居ないだらうといふ意味の事を云つたさうである。此の尊敬すべき大家の謙遜な言葉は今の科學で何事でも分る筈だと考へるやうな迷信者に對する箴言であると同時に、又私の所謂「化物」の存在を許す認容の言葉であるかとも思ふ。もしさうだとすると永い間封じ込められて居た化物共も此れから公然と大手をふつて歩ける事になるのであるが、此れも併し私の疑心暗鬼的の解釋かも知れない。識者の啓蒙を待つばかりである。

(昭和三年十一月下旬)

比較言語學に於ける統計的
研究法の可能性に就て



言語の不思議は早くから自分の頭の中に可也根深い疑問の種を植付けて居たものゝやうである。六七歳の頃、始めて従兄から英語の手ほどきを教へられた時に、最初に出逢つたセンテンスは、たしか「猿が手を持つ」といふのであつた。其時、先づ冠詞といふものゝ「存在理由」が甚しく不可解なものに思はれた。The (當時假名で書く通りにジと發音して居た) が、到る處文章の始め毎に繰返されて出現する事が奇妙に強い印象を與へた事を記憶する。自分の手のことを「持つ」といふものをかしかつたが、此れが「手を」の前に來る事が甚しく不思議であつた。

今になつて考へて見ると、此のジ、ジといふ音の繰返しは、當時の幼い頭の中に、未だ夢にも知らなかつた、遠い／＼所にある、一つの別な珍らしい世界からのかすかなおとづれの

やうに響いたのかも知れない。それは兎に角、當時に感じた漠然たる不思議の感じは、年を経て外國語に對する知識の増すと共に、次第に増しはしても、決して減りはしなかつた。唯それが次第に具體的な疑問の形をとつて意識されて來たのである。併し四十餘年前に漠然と感ぜられた疑問は今日に到つても依然たる不可解の疑問である。そして少し計り言語に關する學者の所説などを讀んで見ても、中々簡単に此の疑問の答解は得られさうもないやうに思はれた。

英語や獨逸語と段々に教はるうちに、屢々日本語とよく似た音をもつた同義の語に出逢ふ事がある。此れは偶然であらうとは思つても、その悉くが偶然の暗合であるといふ事を證明する事も可也に六かしさうに思はれた。

自分のまだ學生時代に、或る學者が、日本の神話の舞臺をギリシア近邊へ持つて行かうとする大膽な説を公にして問題になつた事がある。自分は直接に其所説の全部を讀んだ譯ではなかつたが、其説の一部を何處かで瞥見して、色々其所説に對する疑を起した事もあつた。併し單に説の奇矯であり、常識的に考へてありさうもないといふだけの理由から、此の説を初めから問題ともしないで徒らに嘲笑の的にしようとする人のみ多い事にも疑を懐かない譯には行か

かつた。少くも東歐の一部と極東日本との間に萬一存在したかも知れない何等かの古い關係の可能性といふ事迄も、何の考察もなしに否定せんとする人の餘りに多いのに驚いた。勿論當時此れに關する言語學者間の意見が如何なるものであつたか自分は知らない。こゝで自分のいふのは、言語學者でない一般有識階級と稱するものに就てある。兎に角ギリシア古代と日本古代との間に何等の交渉もなかつたといふ事を科學的に證明する事を果して誰れが敢てし得るであらうと疑つたこともある。

十年程前に少しばかり露西亞語の初歩を學んだ事もあつた。それが爲に「言語の不思議」に對する自分の好奇心と疑問とは、寧ろ急に大きな高い階級を一つ駆け上つたやうな氣がした。そして、一方で新しい不思議が多量に加はると同時に、他方では此の新しい不思議が、却つて古い不思議の謎を解く鍵となり得る可能性を暗示するやうにも見えた。それは單に語彙中の或るものゝみならず、其の文法や措辭法に、東西を結付ける連鎖のやうなものを認める、と思つたからである。

最近に到つて「言語」に對する自分の好奇心を急激な加速度で増長せしめるに至つた経路は

或は一部の讀者に興味があるかも知れないし、又自分が本分を忘れて、他人の門戸を覗ふやうな不倫を敢てするに到つた事の申譯にも幾分はなるかも知れないから一つの懺悔話として茲に誌して見よう。

地球物理學上の近年の問題となつて居る陸塊の水平移動に關する學說、俗に大陸漂移論と稱するものから見た日本陸地の成立、變化、並に此れに聯關して問題となるべき陸地の昇降、地震、火山現象等を追究するに當つて、屢々古い過去に於ける水陸分布の状態と現在のそれとの異同が問題となり、その一つの參考資料として色々な土地の地名の意義が引合に出る場合がある。そこで本邦地名の問題に觸れるとなれば、自然の勢で、アイヌ語や朝鮮語に據る地名起原説を参照しなければならぬ事になる。さうなると問題は自然々に推移して結局は日本語の成立問題に迄も多少は觸れない譯には行かなくなるのである。

さうかと云つて、自分で此のやうな問題をどうにかしようといふ非望を企てる譯にも行かない譯であるが、それでも唯止み難い好奇心から、餘暇ある毎に少しづつ、段々に手近い隣接國民の語彙を瞥見する事になり、それが次第々々に西漸して所謂近東から東歐方面迄も、極めて

皮相的ながら覗いて見るやうな行きがかりになつて來たのである。

かういふ素人遊戯——自分では眞劍なつもりであつても、専門の學者の立場から見れば結局かういふより外はないであらう——に耽つて居る一方で、豫てから、此れも道樂として、心がけて居る日本樂器の沿革に關する考證は自然に世界各地の樂器の比較に移つて行く、其途中で、遠く懸け離れた各種民族の樂器が、其の樂器としての本質に於てのみならず、亦其名稱に於ても、一脈の連鎖によつて互にながつて居るらしく見える現象に逢着して、奇異の感に打たれる事も屢々あつた。勿論樂器の原理は物理學的に普遍なものであるから、各國に同一な樂器のあるのは當然であり、又樂器の名稱が往々擬音から生ずるとすれば、類似の名稱のあるのは當然であると云つて、簡単に片付けて投出して仕舞へばそれ迄である。併しそれで打切つてしまふのは少し危険であると思はせる理由が色々他の方面から供給されるやうである。

少し唐突ではあるが地球上に於ける蚯蚓の分布を調べた學者の研究の結果によると、或種の蚯蚓は、東は日本から海を越えて大陸に、歐亞大陸を横斷して西は西班牙の果迄擴がり、更に驚くべき事には大西洋を渡つて北米合衆國の東部に迄も分布されて居るのである。大陸移動説

を唱へたウェーゲナーは、此事實をもつて歐羅巴と北米大陸とが往昔連結して居たといふ自説の證據の一つとして此れを引用して居る位である。それは兎に角、あの運動遲鈍なみいずさへ、同じ種族と考へられるものが、「現時の大洋」を越えて迄も擴がつて居るといふ事實を一方に置いて考へて見る。勿論此の蚯蚓の先祖と人間の先祖とどちらが古いかといふやうな問題はあつてもそれは別として、此の事實は兎も角も、過去の世界中の人間の間の相互の交渉は、普通想像されて居るよりも、想像され得るであらうよりも、もう少し自由なものではなかつたかといふ疑を喚起させるには十分であらうと思ふ。

世界中の人間の元祖が一つであらうといふ事は單に確率論的の考察から一番考へ易い事であるが、今此處で輕々しくさういふ大問題に觸れようとは思はない。唯少くも動物學上から見て同種な *Homo Sapiens* としての人間の世界の一部分に於て任意の時代に發生した文化の産物の凡てのものが、時と共に擴散して行くのは、丁度水の中に垂らした一滴のアルコールの擴散して行く過程と、何處か類似したものであらう、といふ想像は、理論上それ程無稽なものではあるまいと思はれる。

昔の詩人ルクレチウスは、物質の原子は恰度アルファベットのやうなもので、種々な言語が有限なアルファベットの組合せによつて生ずる如く、各種の物質が此等原子の各種の組合せによつて生ずると書き残したが、此の考は近世になつて化學式といふものによつていくらか科學的に實現された。今此の考を逆に持つて行くとこんな考も起し得られる。即ち、先づ、言語、國語といふ一つの體系は若干の語根元素から組成されて居ると假定する。次には、此の元素が化合して種々の言語や文章が組成されて居るが、此等の間にはその化合分解の平衡に關する極めて複雑な方則のやうなものがあると想像する。なほ此等の元素は必しも不變なものではなくて、例へば放射射性物質の如く、時と共に自然に崩壊し變遷する可能性を有つものと想像する。それで假りに地球歴史の或る一定の時期に於て、或る特別の地點に於て、特殊の國語が急に發生したと假定すると、それは丁度水中にアルコールの一滴を投じたと同様に四方に向つて擴散を始めるであらうと假想される。即ち其の國語の語根の或る一つだけを取つて考へると、それはアルコールの一分子のやうに、不規則に彼方此方と人から人を傳はつて、迂曲した徑路を取りながらも、兎も角も、統計的には、其の出發點から次第に遠く距れて行くで

あらう。尤も、此際問題を複雑にするのは、物質分子の場合と異なり、言語の一分子は獨立の存在として彷徨するのでなく、その周圍に絶えず影響を與へ、自分と同一なものを發生させて行く點にある。併し一つの分子の通過した位では、恐らく其の徑路への影響は短時間に消滅してしまふであらうと考へ、唯同種の分子が種々の徑路を通つて或る地域に到着し、或る時點に於ける其の密度が相當の大きさに達した場合にのみ、其の地點の國語に固定的の影響を與へるであらうといふ、少し無理であるが、又稍尤らしい假定を許容すれば、問題は或度迄は、矢張物質分子の擴散に類したものである。

此の如き假定の下に、或る分子が時間 t に於て、距離 r と、それより dr だけ大きい距離との間の地帯に達するプロバビリテイは

$$W(r,t)dr = \frac{1}{4\pi Dt} e^{-\frac{r^2}{4Dt}}$$

であり、中心から同時に出發した分子總數が N であれば、此の時點に此の地帯に來るもの數は $NW(r,t)dr$ である。併し此等の分子が放射物質のやうに自然崩壊をするものとすれば、此

數は t に就て指數函数的に減じるので

$$Ne^{-\lambda t} W(r,t)dr$$

であるとすべきであらう。さすれば距離 r に於ける密度は、此れを $4\pi r^2 dr$ で除したもので、此を σ とすれば

$$\sigma(r,t) = \frac{N}{8\pi^2 Dt} e^{-\left(\frac{r^2}{4Dt} + \lambda t\right)}$$

で與へられる。

若し中心から不斷に供給が續けられて居れば、此れを時間 t に對して積分する事になるであらう。又中心が空間的に分布されて存在すれば、更に空間的の積分が必要になる事は勿論である。

此のやうな考を實際の場合に應用して具體的の數量的計算をする事は、今の處、不可能であり、又強ひて此れを遂行しても、其價值は疑はしいものである。併し、唯、以上の考察の中に含まれた根本の考が幾分でも實際の問題に觸れた處があるとすれば、右に擧げた數式によつて

代表された理想的過程の内容とその結果とは、又幾分か實際の言語の擴散過程、並に時間的空間的分存の片影を彷彿させる位のものはあるであらうと思はれる。

若しも此の考がいくらか穩當である事を許容するとすれば、そこから色々な、消極的ではあるが、大事な事柄が想定される。即ち先づ世界中で互に遠く隔つた二つの地點に互に類似した言語が存在し、その中間にはその連鎖らしいものが見付からない場合があつても、それだけでは、それが必ず偶然の暗合であるとは斷定されなくなる。又或る甲地方の古い昔の言語が今でも存し、或は今はその地に消滅して居て、其の隣國民乙の間に現存して居るといふ場合に於ても、それだけでその語が甲から乙に移入されたものと推定する事は出来なくなる。何とならばそれは嘗て甲から乙に移つた事があつたとしても、それが甲と前後して乙でも死滅し、すつと後で丙から乙に移つたかも知れないからである。その外分子論的擴散論に於て云はれるやうな色々な事は云はれるが、之を要するに、一つ／＼の言語の分子を比べるだけでは、それだけでは歴史的の前後は決定し難いといふ消極的な結果になるのである。此れは丁度水中のアルコホル分子を一つ／＼捕へる事が出来たにしても吾々は到底その一つ／＼の經路を判定し難いと

同様である。

併し前の考察から一條の活路が示唆される。それは、約言すれば、同系言語の「統計的密度」の「勾配」(Gradient)によつて、その系の言語の擴散方向を推定するといふ方法である。

前の算式によつて示さるゝ如き理想の場合に於ては、一般に同種分子の密度の勾配は、兎も角も中心に對して放射的である。此れは勿論計算を待たずとも明白な事である。それで若し假りに亞細亞大陸の或る地點から或る種の分子が四方に擴散したとすれば、その系統或は同色の言語要素の密度は多少同心圓形分布の形跡を生じてよい譯である。假令此の要素の等密度線がどの様に變形しようとも、少くも、その密度の傾度最大方向のトラヂェクトリーを追跡して行けば、遂にはその源に到着、或は少くも近づく事が出来さうである。

唯第一に問題となるのは、如何なる標準によつてその所謂同系要素なるものを識別し得るかといふ事である。此れは勿論難問題である。併し幸にして從來の言語學者の努力の結果は、此の方法を漸進近似法 (Method of successive approximation) によつて進めんとする際に先づ試みとして置かるべき第一近似の資料を豊富に供給してくれるのである。

此の識別法を假定すれば、次は密度の統計的計算が問題になる。前記の理想の場合の「密度」が直接如何なる數に相應するかは此れも六かしい問題であるが、少くもその一つの計量として、それぞれ地方の國語中に於ける、問題の語系要素の百分率を取つて見る事も一つの適當な試験的方法であらうと考へられる。そして此れは必しも不可能な事とは考へられない。

勿論語根は言語の凡てではない、語辭構成や措辭法も亦言語の要素として重要である。此れ等を如何にして「分子」に分析するかは可也六かしい問題ではあるが、少くも原理の上からはそれも不可能な事とは思はれないのである。

以上のやうな漠然たる想像——勿論此れは今の處唯一つの想像に過ぎない——に刺戟されて、先づ手近な馬來語の語彙に目を通す事を試みた。さうして此の國語と邦語との類似の甚しいのに驚かされた。自然現象や動植物の名稱などはそれ程でもないが、形容詞と動詞に於て特に著しい類似のあるらしい事を感じた。面白い事には、今日我邦一般に行はれて居る極めて卑俗な言語や、日本各地の方言と肖似する現行馬來語も少くない。又試に古事記を繕いて古い日本語を當つて見ると、例へば其中の歌詞——最も古い語の保存されて居るらしい——に現はれた六

かしい語彙などが、可也に尤もらしく、都合よく馬來語で説明され、又古代神名や人名なども、少くも見かけの上で尤もらしく附會されるものが存外多いのに驚かされた。滑稽な例を擧げれば稗田阿禮の名が「博覽強記の人」の意味にこじつけられたりした。又他の方面で最も自分の周圍の人々を愉快がらせたのは、彼の大江山の「酒顛童子」が「恐ろしき悪魔」と譯されたりするのであつた。此れ程關係の深いやうに吾々素人にさへ思はれるものが、何故に今日迄言語學者によつて高唱されなかつたかと思ひ議であるやうに思はれた。現に或る學者の書には、明に馬來と邦語の關係は大したものではないと書いてある。一方朝鮮語やウラルタイ、チャムモンクメール、オセアニック等の語系との關係に就ての論文は往々吾々の目にも入つたが、正面から馬來との關係を論じて、さうしてそれが一般學界ひいては世人の注意をひく程に到つたものゝあつた事は寡聞にして未だ知らなかつたのである。

朝鮮語との語彙の近似は、何人も懐くべき豫期に反して案外に少いものゝやうである。ウラルタイックとも、少くも語彙の點ではそれ程でない事も論ぜられて居るやうである。併し馬來は此の點に就ては恐らく前二者に劣る事はなさうに思はれたのである。

その後 Van Hinlopen Labberton が一九二五年の亞細亞協會學報に載せた論文を讀んで、自分の素人流の對比がそれ程亂暴なものでなかつた事を知ると同時に、外國の學者の間では此れが可也前から問題になつて居る事を知るに至つた。又、Whyman といふ人の「日本語及日本人の南洋起原説」といふものにも出くはした。そして其中で日本人といふものが甚しく低能な幼稚なものとして取扱はれて居るのに不快を感じると同時に其説がそれ程の名論とも思はれないのを奇妙に思つたりした。

馬來を手始めに、アイヌや、蒙古、支那、臺灣など、當つて見ると、勿論可也關係のありさうな形迹は見えるが常識的に豫期される程に密接とも思はれないのを却つて不思議に思つた。それから、ビルマや、タミール、シンガリスなどから、漸次西に向つて、ペルシア、アラビア、トルコ、エヂプト邊をあさつて見ると、矢張幾らかの關係らしいものが認められると思つた。ハンガリーやセルボクロアチアンからフィンランド迄行つて見ても同様である。

併し段々に此調子で漁つて行くと、おしまひにはギリシア、ラテンは勿論現在行はれて居る西歐諸國の語にも矢張同程度の類似が認められる。又かけ離れたアフリカ邊やアイスランド迄

も網の目を擴張られる事になつてしまふのである。

具體的の例は此の序論に於ては省略する積りであるが、唯自分の意味を明にする爲に、試に若干の例を擧げると、例へば、最も縁の遠さうな英語ですらも、強ひてこじつけようと思へば可也にこじつけられない事はない。即ち

beat	butu
laugh	walahu
flat	filattai
hollow	hola
new	nii
fat	futo
easy	yasasi
clean	kilei
ill	walui
rough	araki
hard	katai
angry	ikari
anchor	ikari
tray	tarai
soot	susu
mattress	musiro
etc.	etc.

此の程度のもの、又もつと駄洒落らしいものなら、まだいくらでもありさうである。此等でも、歴史も何も考へずに、子音轉訛や同化や、字位轉換や、最終子音消失やで何とかかかとか理窟をつければ付くであらうし、又中には實際に因果の連鎖のあるものもあるであらう。

もつと思ひ切つて、例へばアフリカへ飛んで Chikaranga の語彙を當ると、一寸當つただけ

で

象	zhou
魚	hoye
鳥	shiri
咽喉	huro

などが見付かる。「象」の訓キサと似たのには馬來の *gajah* (サンスクリットからとある) があるが、ゾウといったやうなのは随分搜したが中々見當りにくくて、それが、どうであらう、突然こんな意外な所に現はれたのである。「魚」も同様であつた。「鳥」は寧ろアイヌの *ahiri* に近いから妙である。土佐で咽喉を切つて自殺する事を「フロヲハネル」と云ふが、此の「フロ」が偶然出て來たのは随分人を笑はせる。尤も萬一ことによると、此れはアラビアの *halq* 其他同系の語を通じて結局は西歐の *gorge, throat, Hals* などにもつながり、又一方例へばベングリの *gale* などを通して可也東洋にも擴がつて居るのかも知れないと想像される。もつと空想を逞しくすれば邦語のゴロなどいふのも少しは怪しくなる位である。(鳥のアラビア語 *tair*. 咽喉の支那語 *hou lung*)。

かういふ種類のでは例へばたつつけ袴のカルサンといふのが印度邊から來て居るかと思ふと、伊太利にも類似の名が出て來たりするのである。(タミール語 *Kalisan*. イタリア語 *Calzoni*)。併し此等の例を挙げたのは、決して此等の語が邦語と因果的に關係して居るといふ事を證明する爲ではなく、寧ろ唯如何なる任意の二つの國語を取つて比較しても、此の種の類似があり得るものであるといふ事の例として取つたに過ぎない。それで例へば、他方で「魚」や「鳥」の訓が支那語や臺灣語で説明されるとか、されないとかいふ事は、茲では問題にならないのである。

兎も角も自分の皮相的な經驗によると、如何なる國語の語彙の比較でも餘りに面白い「發見」があり過ぎるやうな氣がするので、此れは少し考へ方を變へなければならぬといふ事に氣が付いた。さう思はせるもう一つの根據に、A B 兩國語で互に同じやうな音をもつて居ながら意味の方では明白に何の關係もないといふ例が、又可也に多い。最も滑稽な例を挙げるとフィンランド語では鶴が *halkara* であり、狼が *suksi* である。如何にこじつけ度くても、芬蘭の鳥獸と東京の高襟や、江戸前の鯨とを連結すべき論理の絲は見付からない。併しさうなると同じフ

イノ語の狐が Kethu であり、小船が vene であり、樺が koivu であつても、此等の類似の前者の類似との間の本質的の差を説明すべき據り處が分らなくなるのである。

濱の眞砂の中から櫻貝を拾ふ子供のやうな好奇心の追究を一時中止して、稍冷静に立歸つて考へて見ると、此れは寧ろ何でもない事のやうである、統計數學上の込み入つた理論を持出す程でなくとも、簡単なプロバビリテイの考から、少くも原理の上からは、説明のつく事である、といふ風に考へられて来た。

先づ試に、子音にのみ注目するとする。さうしてAの國語に於ける子音の總數を n とする。次に問題を出来るだけ簡單にする爲にB國語の子音をも此れと同數だとする。更に一番簡單な場合を考へて、各子音がそれ／＼各國語に出現する頻度或は確率が一様で、皆に等しいとすると、 $\frac{1}{n}$ で均一になる。(此れは少し亂暴に見えるかも知れないが、統計的方法では多くの場合近似の一法として許される事である。場合により頻度の著しく小さいものは省略する事もやつて見てよい。) 次に語彙中で子音一つより成るもの、二つ、三つ、四つよりなるものといふ風に分類する。そしてそれ等の各々がA B兩國語に現はれる確率をそれ／＼ $a_1 a_2 a_3 \dots$

$b_1 b_2 \dots$ で示すとする。さすればA語の中 i 個の子音より成るもの、中の或る一つを取つて、それと同義の語がB語でも同じ i 個の子音の同順の排列からなるといふ事の確率は $\frac{1}{n^i}$ であると考へる事が出来る。(無論Aで i が2のものを取る場合、Bで i が2でないものは此統計には入れない事にするのである。) 但し此れはA語一つに對するB語に同じ i 級のシノニムが他にないと假定する場合で、もしシノニムがそれ／＼ s_i 個づゝあるとすれば此の確率は s_i 倍に増加する。若し此上にメタセシスを許し、又A語の一子音に對すべきB語子音の轉訛範圍を擴張すれば此れは更に増加する。それが如何に増加するかは計算しようと思へばされる筈のものである。併し茲では最も簡單な場合として、同數シノニムといふ迄に止めると結局A B兩國語彙一般の比較によつて得らるべき純偶然的一致の確率は、

$$P = s_1 a_1 b_1 + s_2 a_2 b_2 + s_3 a_3 b_3 + \dots$$

で與へられる筈である。此の中に出現する s, a, b, v の各數は兎も角も統計的に何とかして求められ得る性質のものである。

以上は出来るだけ事柄を簡單に考へた考へ方である。此れ以上に段々試験的、近似的假定を

修正して、少しづつ實際の場合に近づけて行く事も、原理上からの困難はなく、唯次第に計算が込み入るだけである。併し、今の處、餘りに込み入つた計算では實用にならないから、出来るならば簡単な形で進み度い。

それで第一の試みとしては、先づ前記の一番簡単な場合になるべく適合するやうに、材料の方を選定し排列する事である。それは例へば兩國語の適當な語彙から比較に不適當な分子、例へば本質的でないと思はるゝ接頭語、接尾語などを整理し（勿論此れには或る假定を要するが、それが tentative method として許容される事は、所謂精密科學に於ても同様である。そして此の假定には從來言語學者の苦心研究の結果が全部有効に利用される筈である。）さうしてそれについて上記の a, b を出し、 s は「近似的平均値」を推定して導入する。茲で一番困難なものは A, B の n を同一に整理する事であるが、此れにも色々な方法がある。例へば A 日本語と B 英語の場合ならば、先づ日本語の方を、假りに「日本語ローマ字」で書く、而して英語子音の「文字」の中で日本式にないものは假りに後者のどれかで「置換」する。例へば c や q を皆 k に直す類である。複子音も同様である。 x などは省いても、何かで置換しても統計の結果の値には

大した影響は與へない事は明である。アラビアなどゝなると、大分此の置換が困難な問題となるが、併し例へば喉音の或ものは半數だけ k か g 、残り半數を h で代用するといふやうな試験的便法を取つて第一歩を進める事も出来る。（此處に統計的方法の長所があるとも云はれる。）又例へば馬來語の場合ならば *ber, mer, per* などのプレフィックスの r を省いて見るとか、中間の ng を省いて見るとかする事も試みてよい譯である。

此の如き試験的^{テスタイ}の整理によつて、兎も角も兩國語の子音がそれゝ假りに十四になつたとする。次に假りに $a_1 a_2 a_3 a_4 b_1 b_2 b_3 b_4$ がいづれも $1/4$ で $a_5 b_5$ 以上は零とし、 $s_1 s_2 s_3 s_4$ が平均皆 4 だと假定すると

$$s_i a_i b_i = \frac{1}{4}, \quad i = 1, 2, 3, 4.$$

$$P = \frac{1}{4} \left(\frac{1}{14} + \frac{1}{14^2} + \frac{1}{14^3} + \frac{1}{14^4} \right).$$

$$\frac{1}{14} = 0.07144444$$

$$\frac{1}{14^2} = 0.00510204$$

$$\frac{1}{14^3} = 0.00036443$$

$$\frac{1}{14^4} = 0.00002603$$

$$P = 0.07693691 \div 4 = 0.0192$$

即ち、指定の如き比較によつて、全然偶然から来る暗合の率が約二プロセントは出来る事になる。

併し、上の假定で明に最も不都合なのは、子音唯一つをもつ語の割合を甚しく大きく見過ぎた事である。此れは支那語の場合の外には明に適用されない。

それで、假りに、單子音語の確率を著しく小さいとして度外視し、なほ次の如く假定する。

$$a_1 = b_1 = 0; \quad a_2 = a_3 = b_2 = b_3 = \frac{4}{10}; \quad a_4 = b_4 = \frac{2}{10}.$$

$$\therefore P = 4 \left(0.16 \times \frac{1}{14^2} + 0.16 \times \frac{1}{14^3} + 0.04 \times \frac{1}{14^4} \right)$$

$$= 4 \times 0.0003756 = 0.0035$$

即ち僅に〇、四プロセント弱位に減じてしまふのである。

なほ、若しも、シノニムの數が、上記4の二倍であるとすれば、以上の百分値は矢張二倍になるだけであるから、此の方から結果の桁數オビに著しい影響は起らない。

次に特別な場合として、邦語を假名一つ／＼に切り離し、其の一つ／＼と音韻の似た原語と同義の支那文字を求め、其れを接合して説明をするといふ、普通よくある遣り方をするとどうなるか。此の場合は、 $a_1 b_1$ いづれも1で他は零となるから

$$P = s_1 a_1 b_1 \frac{1}{14} = s_1 \times 0.0714$$

然るに支那では異音類義の字が多いから此の s_1 が大きくなり得る。假りに s_1 を5とすると、三五、七プロセントといふ多數の暗合を見る事になる。此れは此の種の方法による比較の價値を判断する際に参考になると思ふ。なほ此の場合に同じ漢字の發音に對して、各地方的發音の異なるものを材料として、其中から都合のいゝものを探るとなると s_1 が更に一層甚しく大きくなつて、結局どうでもなるといふ事になり、此の如き比較の言語學上の價値は極めて稀薄になつて來る事は明である。

次に比較の標準を少し下げて、メタセシスを許容すると、 P の展開式の i 項に i が乗せら

れる事になるが(但し子音が皆異なるとして)、此れでは少くも n が餘り小さくない限り、明に最後の結果の桁數に變化は起らない。

次に、子音轉訛を擴張して行くと、上記の n が減少し、 r が増加するから、此れは P に重大な影響を及ぼす事となる。假りに濁音を清音と同じにしたり、 k と h 、 m と b 、 s と t などを同一視したりいろ／＼して行くと、獨立したものの數 n は僅々五つか六つになつてしまふ。従つて最後の P は著しく増大する。例へば、 n を 5 とすると

$$\frac{1}{5} = 0.2; \quad \frac{1}{5^2} = 0.04; \quad \frac{1}{5^3} = 0.008; \quad \frac{1}{5^4} = 0.0016.$$

であるから $a_1 = b_1 = 0; a_2 = b_2 = a_3 = b_3 = \frac{4}{10}; a_4 = b_4 = \frac{2}{10}$ の場合 $P = s \times 0.007744$ となり、 s が 4 ならば、約三、一%を得る譯である。即ち、三分位の符合では偶然だが、偶然でないか分らない事になる。

以上は勿論可也色々な無理な假定の下に行つた計算である。此れを逐次修正して言語學者の要求に應ずるやうに近づけて行くことは必しも困難ではないが、茲ではしばらく此れ以上に立ち入らない事にする。

要するに此れは、表題にも掲げた通り、比較言語學上に於ける統計學的研究の可能性を暗示する爲の一つの試みに過ぎないのである。

學者の中には、二つの國語の間の少數な語彙の近似から、大膽に二つのものゝ因果関係を歸納せんとする人もあるやうであり、又一方に於て餘りに細心で潔癖な爲に、暗合の惡戯に欺かれる事を恐れて此種の比較に面迫することを回避する人もあるかも知れない。自分には此の二つの態度がいつまでも互に別々に離れて相對して居るといふことが斯學の進歩に有利であらうとは思はれない。寧ろ進んで、暗合的なものと因果的なものを含めた全體のものを取つて、何かの合理的な篩にかけて偶然的なものとは必然的なものとを篩ひ分ける事に努力した方が有利ではあるまいか。さうして統計的に期待さるべき暗合の確率と、實際の統計的符合率とを對照して、因果關係の「濃度」を示すべき數値を定め、其の値の比較的大なるものに就て、更に最初の假定の再吟味を遂行し、其結果に基いて修正された新たな假定を設け、逐次此の如くして所謂漸近的近似法によつて進行すれば、少くも現在よりは、いくらか科學的に研究を進められはしないかと考へるのである。

例へば子音轉訛の方則の如きでも、獨斷的の考を捨て、可能なるもの、中から甲乙丙……等の作業假定を設けて、此等にそれ／＼相當するPを算出し、又一方此の假定による實際の比較統計の符合の率を算出し、此の兩者を比較して、その結果から甲乙丙いづれが最も穩當であるかを決定すべきである。

統計的方法の長所は、初めから偶然を認容してかゝる點にある。色々な「間違ひ」や「杜撰」でさへも、最後の結果の桁數には影響しないといふ處にある。そして、關係要素の數が多くて、それ等相互の交渉が複雑であればある程、却つて此方法の妥當性がよくなるといふ點である。

それで、此方法を眞に有効ならしむるには、寧ろあらゆる獨斷、偏見、臆説をも初めから排除する事なく、なるべくちがつたものを悉く一先づ取り入れて、凡ての可能性を一つ／＼吟味しなければならぬ。輕々しい否定は早急な肯定よりも遙に有害であるからである。此れは實驗的科學を研究するものに周知の事である。又往々にして忘却される事である。尤もかういふ丹念の吟味をするには可也の手數と時間を要する。それかと云つて、いつ迄も何等か此種の方法をとらなければ、獨斷と獨斷との間の討論の終結する見込は立たないやうに思はれるのである。

如何に面倒でも遂行すればするだけ、後戻りはしないであらうと信ずる。しかも其方専門の研究者の専門の仕事として見るときは、他の科學者、例へば天文學者、物理學者、化學者などの仕事に比してそれ程に面倒な仕事とは決して思はれないのである。

勿論、此れも他の科學の場合と全く同様に、初めからさううまくは行かないであらう。さうして、凡ての可能なるものへの試みの「不可能」を「證明」し、抹殺する事にのみ興味をもつ「批評家」の批評を受けなければならぬ。併しあらゆる「精密科學」はその根柢に於て、丁度此の如き方法を取つて進んで來たものである。凡てが其のはじめは不精密なる經驗の試験的整理を幾重となく折返し繰返し重ねて、漸進的に進んで來たものである。その昔、獨斷と畏怖とが對峙して居た間は今日の「科學」は存在しなかつた。「自然」を實驗室内に捕へ來つてあらゆる稚拙な「試み」を「實驗」の試練にかけて篩ひ分けるといふ事、その判斷の標準に「數値」を用ゐるといふ事によつて、はじめて今日の科學が曙光を現はしたと思はれる。もし古來の科學者が、「試み」なしの臆斷を續けたり、「試み」の結果を判斷する合理的の標準なしに任意の結論を試みたり、或は「試み」に伴ふ怪我のチャンスを恐れて、誰も手を下す事を敢てしな

かつたら、現在の吾々の自然界に關する知識と利用收穫は依然として復興期以前の状態で足踏みをして居たであらう。そして又現在の進歩した時代から見た時に幼稚に不完全に見えないものが如何なる初期の科學の部門に見出され得るであらうか。

餘談は姑く措いて、 AB 、 AC 、 AD ……の關係、なほ念の爲に比較の主客を置換して BA 、 CA 、 DA ……の關係の濃度に對する大體の比較的の數値を定める事が出来たとすれば、少くも茲に AB といふ一つの「鏈の環」が、從來よりは稍科學的な根據の上に假設される。さすれば次には、前に A に就て行つたと同様の方法を、今度は B に就て行ふべきである。さうして兎も角も、 BC といふ、「次の環」の見當を付ける。順次此の如くして、出来るならば又、世界の各方面から出發して、同じやうにして、それ〴〵の鏈を——勿論さういふ鏈が存在するとの作業假説の下に——手繰つて行く。若し多くの人の信するであらう如く、此の數々の鏈が世界の何處かに自然と集合すれば簡單である。さすればその焦點に集中した要素を稍確かに把握し得らるゝから、今度は逆の順序によつて此の焦點から發散し擴散した要素の各時代に於ける空間的分布を驗する事が出来る、その時に至つてはじめて、此の篇の初めに出した擴散に關する數

式が稍具體的の意義を有つて現はれて來るであらう。尤もそれは出来るとしても甚だ遠い未來に於て始めて實現され得る事であらう。

併し上に考へた鏈は恐らく一點には集中しないであらう、それがどう喰ひ違ふか、其處に最も興味ある將來の問題の神秘の殿堂の扉が遠望される。此の殿堂への一つの細道、その扉を開くべき一つの鍵の、臆氣な、しかも拙な詞で表現された暗示としてのみ、此の一篇の正當な存在の意義を認容される事が出来れば著者として寧ろ望外の幸である。

自分は出来るだけ根據なき臆斷と推理を無視する空想を避けたつもりである。併し行文の間に少しでも臆斷の匂があればそれは不文の結果である。推理の誤謬や不備があればそれは不敏の致す所である。此の甚だ僭越と考へらるべき門外漢の一私案が、若し専門學者にとつて何等かの參考ともならば、著者としての欣びは此れに過ぎるものはない。

想ふに此の私案の第一歩の試みを最も有効に遂行する爲には、恐らく言語學者と科學者との協力が需要ではないかと思はれる。若し此の兩者が共同し、その上に器械的の計算や統計を擔當する助手の數人の力をかりることが出来れば、仕事は可也面白く進行しさうに思はれる。併

し此の方が寧ろ恐らく夢のやうな空想であるかも知れない。

(附記) 以上の考察に於ては、最も此種の取扱に便宜だといふだけの理由から、單に「語彙」「單語」のみを問題として、語辭構成法や文法上の問題には少しも觸れなかつた。併し自分は決して後者の比較の重要な事を無視して居るのではない事を斷つて置き度い。尤も文法の如きものでも、此れを數理的の問題として取扱ふ事が必しも不可能とは思はれない。事柄が、見方によつては或る有限數の型式的要素の空間的排列の方式に關するものであると見る事が出来るからである。輓近の數學の種々な方面の異常な進歩は寧ろ色々な新しい此方面の應用を暗示するやうである。又「除外例」といふものゝある事から起る困難は、統計的方法の利器によつて、少くも或度迄救はれ得る見込がある。此れに就ては、更に、機會があつたら、幾分具體的に考を進めて見たいといふ希望をもつて居る。

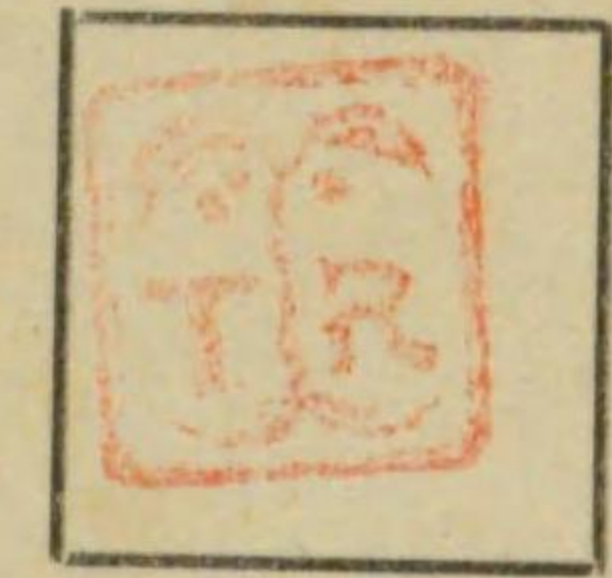
最後に誤解のない爲に斷つて置く必要のあるのは、從來とても統計的のやり方があるにはあるが、單に數をかぞへて多いとか少いかいふだけでは何等の本當の統計としての意味がないといふ事である。全體に對する實際の符合率が偶然による符合率に對する比のみが意味をもつ、茲ではそれを問題にしたといふ事である。

(昭和三年二月)

58
25

昭和四年四月十五日印
昭和四年四月十八日發行

【定價金壹圓八拾錢】



著者 寺田寅彦

發行者 小林立勇
東京市神田區一ツ橋通九

印刷者 白井赫太郎
東京市神田區錦町三ノ一七

發行所

東京市神田區一ツ橋通九番地
振替東京一三七八九番

鐵塔書院

【刷印社興精】

58
25



586
253

