

フィンランド語の名詞の意味と場所格の使用頻度の 関係について

—コーパスのデータに基づく研究—

松村 一登*

1. カテゴリーとしての場所格の使用頻度

本稿で、場所格(local cases)と呼ぶのは、フィンランド語の名詞の活用形のうち、場所関係を表すとされる次の6つの形式のことである。

表1 フィンランド語の場所格の體系

	内部場所格 「内部・密着」	外部場所格 「表面・近傍」
位置格「～で」	内格(inessive)	接格(adessive)
起點格「～から」	出格(elative)	奪格(ablative)
着點格「～へ」	入格(illative)	向格(ablative)

* 東京大学大学院人文社会系研究科

フィンランド語の標準的な文法記述¹⁾に従えば、内部場所格(interior local cases; 以下「内部格」)は、周囲を二次元的に限定された領域、ないし、周囲を三次元的に限定された空間の内部にあるか、平面・物体の表面に密着した状態を表す。他方、外部場所格(exterior local cases; 以下「外部格」)は、二次元的な平面の表面の上、または三次元的な物体の外側・近傍を表す。それゆえ、内部格は「～の中」、外部格は「～の上」ないしは「～のところ」のような日本語の表現に対応していると見なせる場合が多い。

位置格(location)は、日本語の助詞では「～にて、～で」、起点格(source)は「～から、～より」、着点格(goal)は「～へ」にほぼ相当する。英語の前置詞の場合は、内部格を in/out of/into, 外部格を on/(off) from/(on)to に対応させて理解するのが慣例的である。

フィンランド語(書きことば)の名詞の格の使用頻度に関しては、すでにくつかの統計が明らかにされているが、標準語を扱ったものとしては Karlsson (1983:308) [以下Kと略す] と Leino (1993:176) [以下Lと略す] がよく知られている。ともに、それぞれの格を1つのカテゴリーと見なし、比較的小さなサンプル(K - 20,000語, L - 30,000語)に基づき、カテゴリーとしての出現頻度を数えている。この2つの統計から、場所格に関する数字だけを抜き出したのが、表2と表3である。

表2 Karlsson (1983)

	内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)	合計
A	993	613	871	610	145	315	3547
B	7.1	4.4	6.3	4.4	1.0	2.3	25.5

1) 英語で書かれたスタンダードな文法書として Karlsson (1999)に場所格の用法の比較的詳しい説明がある(pp.107-122)。日本語の簡単な解説は、『言語学大辞典』の項目「フィンランド語」(第3巻)、および項目「格」(第5巻)の「フィンランド語の格」にある。フィンランド語の名詞の格に関する詳細な記述は、Penttilä (1963:371-434)を参照。6つの格について、本稿では、それぞれ「内(ine)」「出(ela)」「入(ill)」「接(ade)」「奪(abl)」「向(all)」と略す。

- A. それぞれの格のサンプル全体における出現度数
- B. 名詞の総出現度数においてAが占める割合(%)

表3 Leino (1993)

	内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)	合計
A	27	17	20	13	4	10	91
B	8.0	5.0	5.8	3.7	1.1	2.8	27.2

- A. それぞれの格のサンプル1000語あたりの出現度数
- B. 名詞の総出現度数においてAが占める割合(%)

表2 と表3 から、次のようなことが明らかになる。

- (1) 書きことばにおいて、名詞類(名詞, 代名詞, 形容詞, 数詞など, 格変化を行う語)の全体の出現度数のおよそ4分の1 (K - 25.5%; L - 27.2%)が場所格形である。
- (2) 内部格は、全体として外部格より出現頻度が高い。内部格の中で最も出現頻度の低い出格の頻度は、外部格の中で最も出現頻度の高い接格の頻度よりいくぶん高い。
- (3) 内部格, 外部格を問わず, 出現頻度は, 位置格が最も高く, 起点格が最も低い。
 内格(K - 7.1%; L - 8.0%), 入格(K - 6.3%; L - 5.8%), 出格(K - 4.4%; L - 5.0%)
 接格(K - 4.4%; L - 3.7%), 向格(K - 2.3%; L - 2.8%), 奪格(K - 1.0%; L - 1.1%)

KとLの統計データは、より大きなサンプルに基づいて検証あるいは修正されるべきであるが、より大きなサンプルに基づいた統計データがないのは、大規模なコーパスが存在しなかったり、存在しても、名詞の場所格ごとの出現頻度を調べる手段がなかったことなど、技術的な制約によるものであると考えられる。

参考として「フィンランド語テキストバンク」(Suomen kielen teksti-pankki, 以下「SKTPコーパス」)²⁾の部分コーパス(雑誌記事；ディレクトリ：parole-fi/aikakauslehdet；200万語)に対して、筆者が予備的に行った検索で得られたそれぞれの場所格形の出現度数を表4に示す。

表4 雑誌テキスト(200万語)における場所格形の出現度数

内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)	合計
62367	54907	68650	43884	10117	22682	262607

表4の数字は、内格(内部位置格)と入格(内部着点格)の出現度数の大きさが逆転している点を除けば、総体として表2のKarlsson(1983)、表3のLeino(1993)の提示する統計と基本的な点ではほぼ一致していると思われる(図1参照)。

以下、本稿では、便宜的に Karlsson (1983) に示された数値³⁾が、フィン

2) 「フィンランド語テキストバンク」については、オンライン文書 Miettinen, Manne 2000. Kielipankin asiakkaan opas. CSC - Tieteellinen laskenta Oy (<http://www.csc.fi/kielipankki/opas/x633.htm>)を参照。

3) 表4の統計数字は、SKTPコーパスの雑誌記事からなる部分コーパスを、(1)まずフィンランド語の形態分析プログラムFINTWOLによって単語ごと分析し、(2)その分析結果から、6つの場所格形の1つである可能性をもつ形態をすべて抽出し、数えて出したものである。FINTWOLの分析は context-freeであり、その出力に対する disambiguationが必須だが、それを行っていない生の出力であって、場所格の出現頻度を表すものとしては最終的なものではない。また、表3は、それぞれのカテゴリーの数値が小さいため、本稿で行っているような処理には不適切と思われる。FINTWOLについては、開発元のLingsoft(Helsinki, Finland)の

ランド語の名詞の格のテキストにおける使用頻度を表すものとして信頼できるとみなして議論すすめる。

2. 個々の名詞の場所格ごとの出現頻度

本稿の目的は、フィンランド語において、テキストにおける個々の名詞の場所格ごとの使用頻度は、カテゴリーとしての場所格の使用頻度とは別個の事象であり、それ独自に解明される必要があるということ、コーパスを利用して得られたデータに基づいて実証的に示すことである。なお、筆者の知る限り、フィンランド語の名詞の格の使われ方について、このような観点から正面から論じた研究はみあたらないこと、また、統計学的なデータ処理そのものが筆者にとって新しい試みであること、この二つの意味で、本稿は、今後の研究の方向の模索のための準備段階のレポートとして位置づけられる。

フィンランド語の6つの場所格をカテゴリーと考えたときの、それぞれの場所格の形式がテキストの中で使用される頻度(表2)⁴⁾と、本稿におけるように、特定の名詞がテキストの中で用いられる頻度を場所格ごとに見たものとの間に直接的かつ密接な連関があるならば、どの名詞をとっても、

WWW ホームページ参照。cf. <http://www.lingsoft.fi/doc/fintwol/>

4) 表2～表4のような統計を、フィンランド語の名詞の格の文法的カテゴリーとしての使用頻度を表すものと見なすことができるかどうかについては、議論の余地があろう。ここで、Karlsson (1983) がテキストにおける「格語尾」の使用頻度という言い方をしているのは象徴的である。フィンランド語の場合、形容詞修飾語は主名詞と格と数において呼応するため、名詞句1つにつき、格語尾は「形容詞修飾語の数+1」個出現する。また、形容詞修飾語に関しては、具格や共格のように、1個以上出現することがまれな格形もある。この意味で、形容詞修飾語についた格語尾の数を含めた統計数字を、文法カテゴリーとしての格の使用頻度の研究に用いることは、必ずしも適切とは言えないであろう。本稿は、このような問題を意識した上で、「格語尾の出現頻度」を「文法的なカテゴリーとしての格の出現頻度」の近似値と見なして議論を進める。

場所格ごとの出現頻度は、それぞれの場所格のテキストにおけるカテゴリーとしての出現頻度とほぼ同じになることが期待される。しかし、これは事実と合わない。

物理的な場所関係を問題にする文脈で、個々の場所格の使われ方が名詞によって著しく異なることが話者にとって自明である場合も少なくない。たとえば、名詞 *metsä*「森」が外部格で出るような文脈を考えることがひじょうにむずかしいこと、逆に、*tori*「広場、市場」が内部格で使われることがまずないことは、文法書などでも、しばしば「言語事実」として言及されている。

ここで、*metsä* と *tori* について、それぞれの場所格形のコーパスにおける出現頻度を、「SKTPコーパス」中の新聞記事の部分コーパス(ディレクトリ : *parole-fi/sanomalehdet*; 総語数1360万語)を使って実際に調べた⁵⁾結果が表5である。この表を見ると、*metsä* と *tori* の場所格での用いられかたに関して、母語話者の内省にもとづく「言語事実」と、コーパスによって実証的に明らかになった実際の使用頻度⁶⁾とがきわめてよく一致しているといえる。

5) 以下、本稿における名詞の場所格の使用頻度に関する数字は、すべてこの新聞記事のコーパスに基づくものである。それぞれの名詞の場所格形は、(1) まずコーパスから正規表現を用いて場所格形の可能性のある形式を含む文をすべて取り出し、(2) 場所格の可能性のある形式をフィンランド語の形態分析プログラム FINTWOL で分析し、(3) その出力から Perl スクリプトによって場所格形と考えられる形式を抜き出し、(4) 最後に *disambiguation* を手で行うことによって抽出し、その数を集計した。なお、技術的な制約のため、表や図では、フィンランド語で用いられるウムラウト文字は、対応するウムラウトのない文字(a,o)に置きかえられている。

6) *metsä* が外部格で用いられた用例の大部分は、「森林で行われる仕事」すなわち「森林作業」「狩猟」を意味する。*tori* の出格の用例は、すべて「AからBになる」という場合の変化の起点を表す。いずれの場合も、物理的な「場所関係」を表す用法から派生した用法と考えられる。

表5 metsä「森林」と tori「広場, 市場」の場所格形の使用頻度(絶対度数)

	内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)
metsä	424	341	303	24	4	36
tori	0	12	3	521	119	208

metsä「森林」がもっぱら内部格で用いられ、tori「広場, 市場」がもっぱら外部格で用いられるという「言語事実」は、「雨が降る」とか「風が吹く」などの意味の動詞に、1人称や2人称の活用形がないといった現象や、「ふつう複数形では用いない名詞」「ふつう複数形のみで用いられる名詞」があることなどと同じ性格の文法的な現象と考えられる。しかし、この現象が「ふつう内部格では用いられない名詞」「ふつう外部格では用いられない名詞」といった形で文法の問題として取り上げられることはあまりなく、辞書に書いておけばいい「語彙的な特異性」「語法の問題」と考えるのが一般的であるという印象を受ける。

しかし、この2つの現象を改めて比較してみたとき、たとえば、英語の動詞 to rain「(雨が)降る」について、一人称主語をとる I will rain tomorrow のような文が、雨が擬人化された童話などのように非常に限定された文脈でも考えないかぎりまず使われないという「(直観的)言語事実」と、フィンランド語の名詞 tori「広場, 市場」の内格形 torissa (「広場の内部で」) がまず使われないという「(直観的)言語事実」との間に、言語現象として決定的な差があることを積極的に示すことは困難である。

ここで、名詞によって場所格ごとの使用頻度に大きな違いがあるということが、言語学的な意味をもっていると考えるべき根拠を2つの観点から整理しておこう。

まず、統計データの解釈という視点から見たとき、本稿でこれから扱うコーパスにおける個々の名詞の場所格ごとの出現頻度と、カテゴリーとし

てのそれぞれの場所格のテキストにおける使用頻度(表2)との間のギャップがきわめて大きいこと、さらに、その際のばらつき方が名詞ごとにまちまちであるという事実(表5)を、単なる偶然として無視することはできない。ちなみに、本稿で考察の対象とした【資料】の名詞57語の中には、場所格形が表2とよく似たパターンで用いられている名詞は見つけることができない。

第2に、場所格の頻度に関する統計的な数値と母語話者の言語直観との間に相関がある可能性が否定できない。表5のように、個々の名詞の場所格ごとの出現頻度が、母語話者の内省的な観察とぴったり符合するケースが見られるという事実を、単なる偶然として片付けるのには無理があり、何らかの言語学的事実の反映と考えるべきである。

以上の理由から、カテゴリーとしての場所格の使用頻度と、個々の名詞の場所格ごとの出現頻度の食い違いには、言語学的な意味があると考えた方が、コーパスから得られた統計的なデータの解釈としてより妥当と考えられる。

カテゴリーとしての名詞の格の出現頻度は、言語学の教科書や論文などに、フィンランド語の格の用法に関する統計的事実としてしばしば引用されることからわかるように、従来、言語学的に意味のある数字であることが自明であるとして扱われてきた。これに対して、個々の名詞の場所格ごとの使用頻度に関しては、先に述べたような直観的観察(「tori は内部格では使いにくい」)を除けば、事実上注意が払われてこなかったという明らかな違いがある。

このような待遇の違いは、理論的な根拠に基づくものというより、大規模コーパスが存在しないなど、決定的ともいえる現実的・技術的な制約があって、個々の名詞について場所格ごとの頻度を調べようにも、現実問題として実行不可能であったという事情に由来すると考えるべきである。カテゴリーとしての場所格の使用頻度と比べたとき、特定の名詞に限ったと

きの場所格の使用頻度はきわめて小さく、数万語程度の規模のテキストで調査することに統計的な意味がなかったため、誰も注意を払おうとしなかったとしても不思議はない。

さらに、仮に大規模なコーパスが存在していたとしても、それが plain text である限り、特定の語の特定の格形のみを抽出することは、形態分析プログラムが開発されるまで、實際上、実行不可能といってよかったという事情がある。大規模コーパスと形態分析プログラムの両方が利用可能になって、新しい可能性が開けたのは、ほんのここ数年のことである⁷⁾。

3. 場所格の使用頻度データの解釈と処理

この節では、表5が示す metsä「森林」と tori「広場、市場」の場所格形の使用パタン間の大きな違いを含め、個々の名詞の場所格での使用パタンが一樣ではない事実に注目し、この事実は、名詞の意味と場所格での使用パタン間に密接な関係があることの反映ではないか、という予測のもとに議論を進める。より具体的には、

【仮説】 意味が類似している名詞は、場所格の使用のパタンも類似している。

という仮説を、コーパスを利用して検証(または反証)することを試みる。まず、この仮説で用いられている概念のうち、「意味の類似」と「場所格の使用パタン」の内容を実証的に検証可能な形に定義しておく必要がある。名詞の「意味の類似」については、認知的なカテゴリーに関する議論におけるように、名詞はシソーラス風に意味が相互に関連した体系を作っている

⁷⁾ 筆者にとって、この2つの条件が満たされたのは、SKTP コーパスへのアクセスの許可が得られた2000年春である。

という理解を出発点として考える。すなわち、「人」「動物」「建物」「交通手段」「季節」「容器」等、名詞の指示対象の性質によって「自然に」きまる名詞のカテゴリの階層的構造を想定した上で、同じカテゴリに属する名詞は互いに意味がきわめて近く、隣接するカテゴリに属する名詞も互いに意味が比較的近い関係にあると考える。また、「意味が近い」ということを「意味の点で距離が近い」とも言うことにする。このように規定したときの「意味的な距離」についての(直観的)判断は、同じ言語共同体に属する話者によって間主観的に共有されていると考える。

「場所格の使用パターン」は、コーパスにおけるその名詞の場所格ごとの出現頻度の相互関係であると考え。具体的には、それぞれの名詞に、コーパスにおける出現頻度を内格・出格・入格・接格・奪格・向格の順で並べた6つの数値の組み合わせを対応させる。たとえば、*metsä* には (424, 341, 303, 24, 4, 36), *tori* には (0, 12, 3, 521, 119, 208) のような6つの数値が対応付けられる。この6つの数字の組み合わせを、その名詞の「場所格の使用パターン」と呼ぶことにする。

2つの名詞について「場所格の使用パターン」の類似性の判定には、大きく2つの方法が考えられる。

【方法1】 6つの数字をプロットしてグラフを描き、グラフの形の重なり具合を視覚的に判断する。グラフの重なりの度合いが大きいほど、2つの単語の「場所格の使用パターン」は類似していると考え。

【方法2】 6つの数字の組み合わせの類似度を計算によって数値として出して、その大小によって、2つの語の「場所格の使用パターン」の類似度を測る。

本稿では、方法1として折れ線グラフ⁸⁾、方法2として階層クラスタ分

析と主成分分析⁹⁾を使ってみた。

2つの方法にはそれぞれ長所と短所があり、相補い合う方法と考えるのが適当と思われる。折れ線グラフは、場所格の使用パターンが近い名詞について、具体的にどのような比率で場所格が用いられるかという特徴を一目で明らかにできるという利点を持つが、2つないし比較的少数の名詞の間での場所格使用パターンの比較でしか効果を発揮しない。これに対して、階層クラスタ分析と主成分分析は、名詞57語の場所格使用パターンを総体としてとらえ、その相互の類似性を一目で鳥瞰することができるという長所があるが、各グループごとの場所格使用パターンの特徴を見るには適していない。

どちらの方法でも、場所格の出現頻度を表す生の数値をそのまま用いるか、それとも、生の数値に一定の方法で補正を施した数値を用いるかを決めておく必要がある。本稿では、以下に述べるような理由で、2つの観点からデータの補正を行ったほうが、生の数値を用いた場合よりの確かな形で場所格の使用パターンの比較ができると判断した。

具体例として、*mies*「男、夫」と *nainen*「女」の場所格の使用パターンを比較する。表6は、2つの名詞の場所格ごとの出現頻度の生の数値で、表6を折れ線グラフにプロットして、2つのグラフを重ねたものが図2である。

8) 折れ線グラフの描画には、RQC(Ver2.44, ©1998-2000 Takeshi Nojima)というフリーウェアを利用した。

9) 階層クラスタ分析と主成分分析は、統計数理研究所の上田澄江さんをお願いした。上田さんには、また、解析結果の解釈のしかたに関してもご教示いただいた。併せてお礼を申し上げる。言語研究におけるクラスタ分析を含む統計学的手法については、Woods et al (1986), Oakes (1998) を参照。類型論に階層クラスタ分析を適用した研究に Tsunoda et al (1995) がある。本稿のための階層クラスタ分析は、「ユークリッド距離、群平均法」「ユークリッド距離、最長距離法」の両方で行われたが、筆者には現時点では解析の結果を本格的に解釈するだけの知識がないので、本稿での扱いは予備的な考察の段階にとどめておくのが適当と考え、問題を複雑にしないために、後者の方法による分析結果の樹形図(デンドログラム)のみを階層クラスタ分析の結果として提示した。なお、解析結果の言語学的解釈は筆者が行ったものであり、データの解釈に問題点がある場合はすべて筆者の責任である。

表6 mies「男, 夫」と nainen「女」の場所格形の使用頻度(絶対度数) cf. 圖2

	内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)	合計
mies	76	714	201	464	241	870	2566
nainen	56	282	103	185	51	374	1051

表6・図2を見ると, mies と nainen の場所格の使用パターンはそれほど類似していないように思われる。しかし, 2つの名詞には, 場所格を総体としてとらえた時の出現頻度の点で, 2.5倍近い開きがある。2つの名詞の場所格における出現度数の合計の大小に起因する「錯覚」をなくすために, 表6の数値を2つの名詞がともに1000回用いられたと仮定したときに予想される出現度数に変換すると, 表7が得られる。この操作を「出現度数の理想化」と呼ぶことにする。表7を折れ線グラフにプロットしたのが図3である。

表6・図2と表7・図3を比較すると, 前者では mies と nainen の場所格における使用パターンの類似が目に見えてこないのに対して, 後者では, グラフがの重なりが視覚的に明確となり, 場所格の使用パターンの類似が明らかになっている。したがって, このような形で「理想化」した数値を用いる方が, 場所格の使用パターンの比較により適していると考えられる。

表7 mies「男, 夫」と nainen「女」の場所格形の使用頻度(表6の理想化) cf. 圖3

	内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)
mies	29.6	278.3	78.3	180.8	93.9	339.0
nainen	53.3	268.3	98.0	176.0	48.5	355.9

表2の場所格のカテゴリーとしての使用頻度を見ると, 格によってテキストにおける頻度にかかなりの開きがあることが明らかである。表2に基づ

き、奪格の出現頻度を1としたときの他の5つの格の出現頻度の比を計算すると、表8のようになり、もっとも頻度の高い内格と最も頻度の低い奪格の比は、6.85対1となっている。

表8 奪格を1としたときの各場所格形の相対的な頻度 cf. 表2

内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)
6.85	4.23	6.01	4.21	1.00	2.17

表8は、それぞれの場所格の使われやすさを相対的に表している。いいかえると、それぞれの格の出現頻度が表8のような比率となる時が、フィンランド語のシステム全体からみたとときのバランスのとれた場所格の使用パターンと考えられる。そこで、場所格が表8のような比率で現れた場合に、すべてのカテゴリーが同一の数字を示すように、表7の数字を補正処理するのが適当と考える。この補正は、もともと内部格と比べると頻度の低い外部格の頻度が、表8の比率と比べて高い名詞において、その事実を際立たせる役割を果たす。

このような観点から表7の数値に補正を施すと表9が得られ、表9を折れ線グラフにプロットすると図4が得られる。

表9 mies「男、夫」と nainen「女」の場所格形の使用頻度(表6の理想化+補正)

cf. 図4

	内(ine)	出(ela)	入(ill)	接(ade)	奪(abl)	向(all)
mies	11.50	174.98	34.67	114.27	249.68	414.91
nainen	22.77	185.72	47.74	122.44	142.00	479.33

表7・図3と表9・図4を比較すると、後者では、miesとnainenが外

部格で使用される割合が、場所格の総体としての使用パターンと比較してかなり高いことが目に見えるが、補正を施していない前者では、もともと標準で向格のほぼ2倍の頻度をもっている出格の頻度の高さが依然として目につくため、この2つの名詞の向格の頻度が標準(表8)と比べてかなり高いという事実が、それほど目立たない。

以上のような考察に基づき、以下では、それぞれの場所格の出現度数の生の値に対して「理想化」と「補正」を行った数値を「場所格の使用パターン」と見なして議論を行うことにする。

4. 名詞ごとの場所格の使用頻度の分析結果の考察

本稿のために、SKTPコーパスにおける場所格の使用パターンを調査した名詞57語の一覧を、コーパスにおけるそれぞれの語の場所格ごとの出現度数とともに【資料】に掲げる。それぞれの名詞につけられている番号は、図5の数字に対応する。

【資料】のデータに理想化と補正を行った上で、前節で規定した方法2の主成分分析によって分析、場所格の使用パターンにおける名詞間の距離の遠近を2次元座標上にプロットしたものが図5であり、同じく、階層クラスタ分析によって分析し、場所格の使用パターンにおける名詞間の距離の遠近を樹形図(デンドログラム)として表現したものが図6である。

名詞間の距離の大小にのみ注目する方法2に対して、距離が近いと判定された名詞の場所格の使用パターンに共通する特徴を明示的に見るためには、図7～図22のような折れ線グラフの助けを借りる必要がある(方法1)。以下、図5、図6に表された結果に図7～図22を照らし合わせながら、議論を進める。

表10 主成分分析で確認できる意味の類似した名詞のグループ

A	季節と1日のうちの時間区分(図7, 図8)
	5 kevat 春 48 syksy 秋 50 talvi 冬 1 aamu 朝 13 kesa 夏 57 yo 夜 7 ilta 夕 37 paiva 日中
B	曜日(図9, 図10)
	39 perjantai 金曜日 53 torstai 曜日 51 tiistai 火曜日 30 maanantai 月曜日 14 keskiviikko 水曜日 27 lauantai 土曜日 47 sunnuntai 日曜日
	家族とその周辺(図11, 図12)
C	34 mies 男 35 nainen 女 20 kissa 猫 25 lapsi 子ども 21 koira 犬 38 perhe 家族
	部屋, 建物(図13, 図14)
D	45 sauna サウナ 5 huone 部屋 4 hotelli ホテル 18 kirjasto 図書館 22 koulu 学校 49 talo 家 19 kirkko 教会

図5では、名詞は【資料】における番号で表されているが、表10の示すように、同一カテゴリあるいは隣接カテゴリに属すると考えられる名詞が比較的近い位置に集まっているケース(図5の太い曲線で囲った単語グループ)がいくつか観察される。

また、A~Dほど大きくないが、何らかの点で共通点があると考えられる2~3語が隣接するか、比較的近くに位置する例もいくつか見られる(図15~図20)。

表11 主成分分析で隣り合わせ、ないし比較的近くに位置する意味の類似した名詞の小グループ

E	12 katu 通り — 52 tori 廣場
F	26 lattia 床 — 40 piha 中庭
G	3 bussi バス — 11 juna 列車 — 2 auto 自動車

H	23 laatikko 箱	—	28 laukku かばん
I	33 metsa 森林	—	43 puisto 公園
J	36 ovi ドア	—	46 seinä 壁

これに対して、図6は隅々までバイナリの枝分かれによって名詞間の距離を視覚化してくれ、大きなグループ分けに対しても、その言語学的意味が比較的是っきりと見える。枝分かれを樹形図の根元に近い方から見ていくと、まず、他の名詞から突出したグループを作っている8つの名詞(グループ1)は、上でグループAとした「季節と1日のうちの時間区分」を表す8語と一致する。実際、グループAの8語は、接格だけが単独で突出して用いられる(cf.図7～8)という点できわめて得意な場所格使用パターンを示すことが確認できる。

次の大きな枝分かれは、外部格が多く使われる名詞と内部格が多く使われる名詞を分けていると考えられる。この大きな区分を便宜的に「外部格型」「内部格型」と呼び分けることにする。外部格型の名詞グループの中では、k_{atu}「通り」や tori「広場」のように、内部格がほとんど用いられない名詞(cf.図15)がグループAに近いところに位置しており、内部格型の名詞グループでは、逆に、metsä「森林」や sauna「サウナ」に代表される外部格の頻度が非常に低い名詞(cf.図21～22)がグループAからもっとも遠いところに位置している。

外部格型の名詞のグループ内の最初の枝分かれが作る2つのグループ(グループ2とグループ3)については、名詞のカテゴリーとしての特徴付けはむずかしい。ただし、グループ3の中の最初の枝分かれが作る2つの下位グループのうちの一つは、上でグループC(「家族とその周辺」)にまとめた6つの名詞のグループとだいたい一致する。この下位グループはさらに2つのグループに分かれているが、nainen「女」kissa「猫」lapsi「子ども」と mies「男、夫」koira「犬」が別々のグループに属しているのは興味深い。ただし、ここにグループCとはおよそ無関係と思われる mäki「山、丘」と maito「乳」が紛れ

込んでいる。これは、カテゴリーが近いほど場所格使用パターンが似てくる傾向があるとしても、パタンの類似は必ずしもカテゴリーの類似を意味しないということを示していると考えられる。

内部格型の名詞のグループ内の最初の枝分かれが作るグループの1つ(グループ4)は、上でグループB(「曜日」)と呼んだグループと比較的うまく重なる。グループ4のきわだった特徴は、内部格の中でも出格と入格の頻度が目立つことであろう。

内部格型の名詞の残りは、外部格の使用頻度が比較的高いタイプ(グループ5)とそうでないグループ(グループ6)に下位分割される。この2つのグループも、グループ全体を特徴づけるような名詞カテゴリーを指摘するのが難しい¹⁰⁾。

こうして得られる6つの名詞グループを整理すると表12のようになる。

表12 階層クラスタ分析に基づく名詞57語のグループ分け

1	aamu「朝」 talvi「冬」 kevat「春」 syksy「秋」 paiva「日」 ilta「夕」 kesa「夏」 yo「夜」
2	katu「通り」 tori「廣場」 jarvi「湖」 ovi「ドア」 joki「川」 tuoli「イス」
3	poyta「机」 seinä「壁」 lattia「床」 piha「庭」 lapsi「子ども」 kissa「猫」 nainen「女」 maki「丘」 perhe「家族」 koira「犬」 mies「男、夫」
4	keskiviikko「水曜日」 perjantai「金曜日」 lauantai「土曜日」 tiistai「火曜日」 torstai「木曜日」 maito「乳」 ikkuna「窓」 maanantai「月曜日」
5	bussi「バス」 juna「列車」 juhla「祭典」 puhelin「電話機」 kieli「舌」 言語」 auto「自動車」 pullo「ビン」 lumi「雪」 laiva「船」 vesi「水」

10) グループ6の中で sunnuntai「日曜日」は特出しており、1語だけのグループを作っていると見たほうがいいかもしれない。この語が抜けると、グループ6は、少しだけ等質性を増すと考えられる。この語は、主成分分析と階層クラスタ分析の結果が大きく食い違うもっとも顕著な例である。

6	kirkko「教會」 koulu「學校」 talo「家」 verkko「網」 laatikko「箱」 laukku「かばん」 puisto「公園」 hotelli「ホテル」 huone「部屋」 kirja「本」 metsä「森林」 kirjasto「圖書館」 sauna「サウナ」 sunnuntai「日曜日」
---	--

なお、図6で2～3語で小さなまとまりを作っているとして表11にまとめた6グループのうち、E, F, G(autoを除く), Hが図7でも枝の先に並んで現れており、場所格の使用パターンにおいて最も近接した名詞どうしであることが、クラスタ分析によっても確認される。

他方、同じく表11-Jの ovi「ドア」と seinä「壁」は、表12ではグループ2とグループ3に分かれている。階層クラスタ分析は、この2つの語の場所格の使用パターンがとくに近いとは認定していないわけだが、これはこの2つの名詞の場所格使用パターンの折れ線グラフ(図20)の相違が著しいことと符合する。この例ほど顕著ではないが、同じく、主成分分析とクラスタ分析の結果の食い違いが比較的大きい表11-Gの auto(図17),あるいは、表11-I(図19)の場合、折れ線グラフはクラスタ分析の方を支持して、場所格の使用パターンの類似性の低さを示していることが多いように思われる。

なお、グループ2(外部格優勢)とグループ6(内部格優勢)の名詞の場所格使用パターンが両極端であることは、中央で交差しているだけで、ほとんど左右対称と形容していいほどの対照的な折れ線グラフの形によって視覚的にも明らかになる(図21, 図22)。

5. まとめと今後の課題

名詞の意味(所属するカテゴリー)と場所格の使用パターンとの間には、方法1(折れ線グラフ)と方法2(主成分分析, 階層クラスタ分析)のいずれから

も、密接な関係があると判断すべきケースが数多く認められる。ただし、*auto*「自動車」と *pullo*「ビン」(グループ5)や *kirja*「本」と *metsä*「森林」(グループ6)のように、意味的な距離が離れていると判断される名詞が、場所格の使用パターンにおいてきわめて近い距離に隣接している場合もあり、名詞の категория が場所格の使用パターンが決まる際の唯一の決定的な要因とはいえないようである。

6つの場所格が、物理的な場所関係を表す場合以外でも広く用いられることは周知の通りであるが、本稿で行った頻度調査においては、当該の場所格が、物理的な場所関係を表しているかどうかを区別していない。しかしながら、外部格型の名詞の中で、外部格の用法の大部分が物理的な場所関係を表すと考えられる *katu*「通り」や *tori*「広場」のような名詞(グループ2)と、内部格の用法の大部分が物理的な場所関係を表すと考えられる *metsä*「森林」や *sauna*「サウナ」のような名詞(グループ6)が、場所格の使用パターンにおいて対極にあることを示していると解釈できる結果が、方法1と方法2の両方から得られるのは興味深い。この結果をもうすこし文法研究寄りの観点から解釈すれば、本来場所を表すことが多い名詞は、外部格型と内部格型に分かれ、そして、このような名詞の場所格は、ふつう、物理的な場所関係を表しているというフィンランド語の伝統的な文法記述の追認となる可能性が高い。

しかし、同時に、すべての名詞に6つの場所格を等しく認定する従来の文法記述を改める必要性があることも示唆される。もし、従来の文法記述がフィンランド語の場所格の意味と用法の実際を的確に表しているならば、すべての名詞において、6つの場所格形がおおむね同じように用いられるはずであるし、また、たとえば、すべての名詞の接格形に、場所、時間、道具、... を表す例がまんべんなく観察されてもおかしくないからである。現実はもちろんその正反対である。

筆者は以前、エストニア語の場所格の1つである接格について、場所格

の1つとして分類されているものの、物理的な場所関係を表している用例の割合が比較的少く、頻度の点でもっとも一般的な接格の用法は、筆者が「与格用法」と呼ぶ用例であることを 37,000語ほどのテキストを調査した結果に基づいて指摘したことがある(松村1992a; Matsumura 1994, 1997)。その研究の過程で次第に明らかになってきたことの中に、エストニア語の接格の用例のうち、場所関係を表しているとされるものを詳しく観察すると、多くの場合、接格名詞句の主名詞の指示対象が、もともと物理的な場所であるか、あるいは、物理的な場所関係を表す際に基準点とされることの多い物体(例「ドアのところ」)であること、言い換えると、場所格で現れなくとも「もともと場所を表している」と見なすことのできる名詞であること、同様に、時間を表しているとされる接格の用例の大部分は、「朝」「夏」「月曜日」のような、もともと時間を表す名詞が接格名詞句の主名詞として現れているケースであることなど、一見トートロジーとも見える事実がある。これは、おそらくエストニア語の他の場所格についても言えるだろうし、さらにはフィンランド語の場所格についてもなりたつと考えられる。

このことは、本稿で明らかにした「フィンランド語の名詞の意味と場所格の使用パターンには相関がある」ということと表裏一体の関係にあると言っている。すなわち、物理的な場所を指示対象とする名詞が場所格で用いられると、物理的な場所関係を表すのがふつうであるばかりでなく、その逆、すなわち、物理的な場所関係を表す接格名詞句の主名詞は、多くの場合、物理的な場所を指示対象とする名詞であるということもなりたつことが示唆されるからである。さらには、時間を表す名詞と時間関係を表す場所格の関係、手にとって道具として使うことのできる物を表す名詞と「～を使って」という道具・手段を表す接格との関係など、名詞の意味によって、その場所格形にふつう期待される意味がかなりの程度まで特定できるのではないかと考えられる。

もしこれが正しいとすると、従来の文法記述にみられる「フィンランド

語の接格には、場所を表す用法、時間を表す用法、道具を表す用法、... がある」という記述方法は不正確であって、むしろ「通常、場所を表す名詞の接格形は場所関係を表し、時間を表す名詞の接格形は時間関係を表し、道具として用いられるものを指示対象とする名詞の接格形は道具を表し、...」と書くべきであるということになる。

これを実証的に裏付けるためには、本稿で行ったような、同じ名詞の場所格形を十把一からげにして出現頻度を数えるのではなく、同じ名詞の同じ場所格形が異なった意味で用いられる場合を区別することによって、場所格の使用パターンをよりきめ細かく明らかにしていく作業と併せて、動詞の側から名詞のカテゴリとは無関係に特定の場所格の使用を指定してくる核支配のような現象が、個々の名詞の場所格使用パターンにどの程度影響をもつのかも明らかにする必要がある。これが今後の課題である。

【資料】 SKTP コーパスにおける名詞57語の場所格ごとの出現頻度

	内格	出格	入格	接格	奪格	向格	意味
1 aamu	7	120	79	583	1	3	朝
2 auto	318	361	345	489	23	78	自動車
3 bussit	107	31	91	116	0	9	バス
4 hotellit	187	68	104	9	1	14	ホテル
5 huone	481	128	250	50	9	27	部屋
6 ikkuna	42	186	48	18	1	12	窓
7 ilta	111	99	243	656	3	7	夕
8 järvi	80	99	122	324	92	10	湖
9 joki	9	3	19	13	4	8	川
10 juhla	379	117	325	250	15	83	祭典
11 juna	145	59	96	134	2	9	列車
12 katu	4	41	38	757	135	273	通り

13 kesa	110	207	151	978	27	14	夏
14 keskiviikko	0	29	38	0	0	0	水曜日
15 kevat	8	116	103	1146	8	11	春
16 kieli	92	208	88	342	16	74	舌, 言語
17 kirja	727	488	388	92	7	41	本
18 kirjasto	146	67	95	3	1	23	圖書館
19 kirkko	537	77	181	57	27	74	教會
20 kissa	2	11	4	6	2	18	猫
21 koira	5	57	34	21	11	37	犬
22 koulu	1241	457	622	154	32	195	學校
23 laatikko	60	55	76	18	1	4	箱
24 laiva	45	63	62	143	17	36	船
25 lapsi	21	376	103	181	98	804	子ども
26 lattia	8	10	30	81	11	93	床
27 lauantai	2	33	34	0	3	4	土曜日
28 laukku	19	13	24	3	0	0	かばん
29 lumi	16	30	19	18	4	9	雪
30 maanantai	0	74	30	0	4	3	月曜日
31 maito	17	43	25	6	0	17	乳
32 maki	16	13	17	21	3	24	山, 丘
33 metsa	424	341	303	24	4	36	森林
34 mies	76	714	201	464	241	870	男, 夫
35 nainen	56	282	103	185	51	374	女
36 ovi	32	112	33	115	50	64	ドア
37 paiva	551	365	540	1637	105	131	日, 日中
38 perhe	237	172	194	62	35	216	家族
39 perjantai	0	52	91	0	1	3	金曜日
40 piha	7	2	5	43	8	34	中庭

41 poyta	222	56	313	143	47	233	机, テーブル
42 puhelin	144	51	124	67	3	5	電話機
43 puisto	384	107	217	7	2	13	公園
44 pullo	33	47	38	48	4	8	ビン
45 sauna	79	36	47	6	3	7	サウナ
46 seinä	37	34	171	120	22	98	壁
47 sunnuntai	1	12	73	0	2	7	日曜日
48 syksy	6	182	131	1439	10	7	秋
49 talo	946	349	460	226	50	131	家
50 talvi	16	61	30	409	6	3	冬
51 tiistai	0	51	31	0	1	3	火曜日
52 tori	0	12	3	521	119	208	廣場, 市場
53 torstai	1	57	48	0	1	2	木曜日
54 tuoli	74	31	44	49	12	21	イス
55 verkko	265	93	305	45	7	25	網
56 vesi	285	167	259	426	72	148	水
57 yö	150	90	87	492	7	7	夜

参考文献

- Karlsson, Fred 1983. Suomen kielen äänne- ja muotorakenne. WSOY. [『フィンランド語の音韻と形態』]
- _____ 1999, Finnish. An Essential Grammar. Routledge.
- Leino, Pentti 1993, Polysemia -- kielen moniselitteisyys. Helsingin yliopiston suomen kielen laitos. [『ポリセミー --- 言語の多義的性格』]
- Matsumura, Kazuto 1994. "Is the Estonian adessive really a local case." Journal of Asian and

- African Studies 46/47: 223-235.
- _____ 1997. "The dative use of the adessive case in Estonian: a corpus-based study." In: K. Matsumura & T. Hayashi, eds., *The Dative and Related Phenomena* (ひつじ書房), pp.31-79.
- Oakes, Michael P. 1998, *Statistics for Corpus Linguistics*. Edingburgh University Press.
- Penttilä, Aarni 1963, *Suomen kielioppi*. WSOY. [『フィンランド語文法』]
- Tsunoda, Tasaku, Sumie Ueda & Yoshiaki Itoh 1995, "Adpositions in word-order typology." *Linguistics* 33: 741-761.
- Woods, Anthony, Paul Fletcher & Arthur Hughes 1986, *Statistics in language studies*. Cambridge University Press.
- 松村一登 1992a, 「エストニア語の接格(adessive)について」, 『文化言語学』(三省堂), pp.993-1008.
- _____ 1992b, 「フィンランド語」, 『言語学大辞典』第3卷(三省堂), pp.672-688.
- _____ 1996, 「フィンランド語の格」, 『言語学大辞典』第6卷(三省堂), pp.201-207.

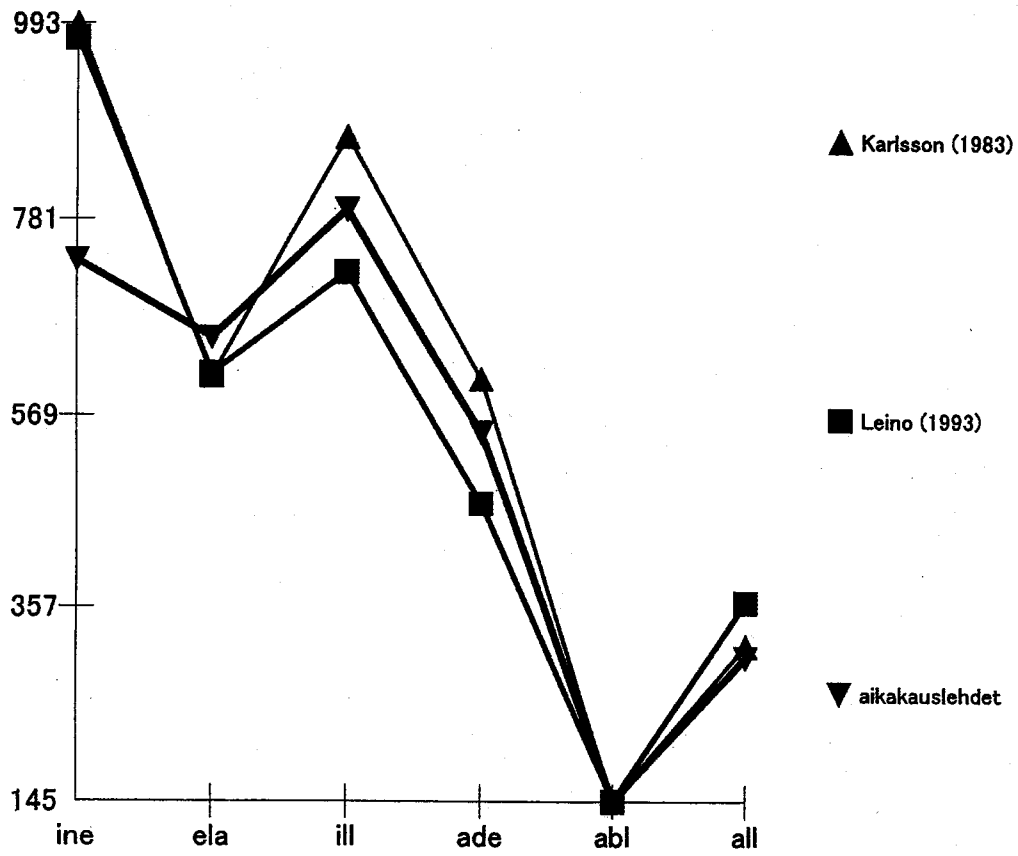


図1 フィンランド語の場所格のテキストにおける出現頻度 (cf. 表2～表4)

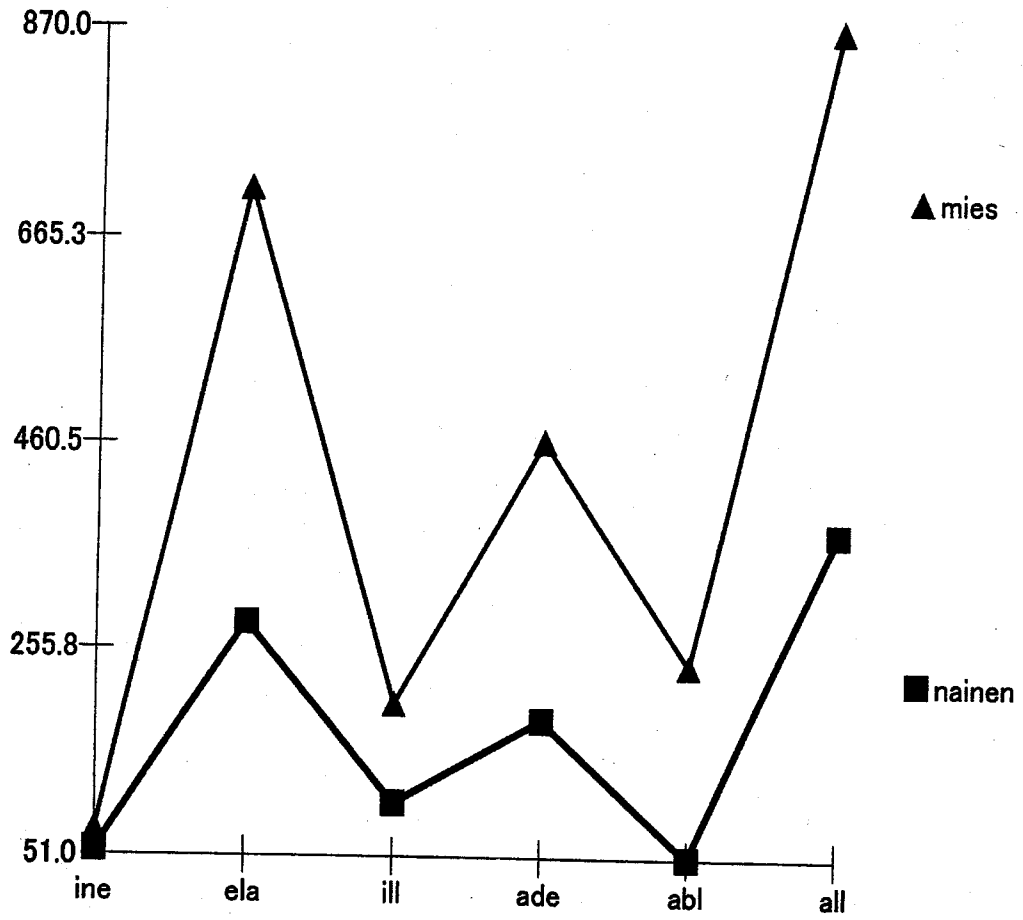


図2 SKTPコーパスにおける mies「男、夫」と nainen「女」の場所格形の出現頻度 (cf. 表6)

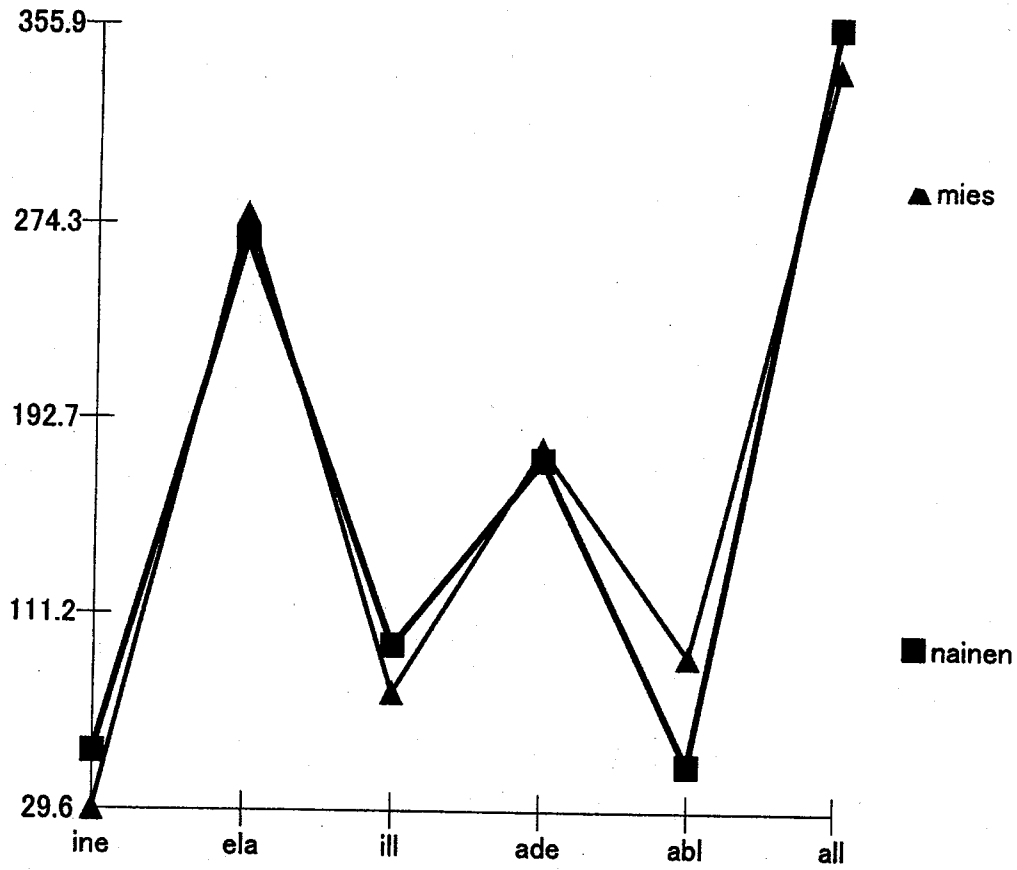


図3 SKTPコーパスにおける mies「男、夫」と nainen「女」の場所格形の出現頻度（理想化）（cf. 表7）

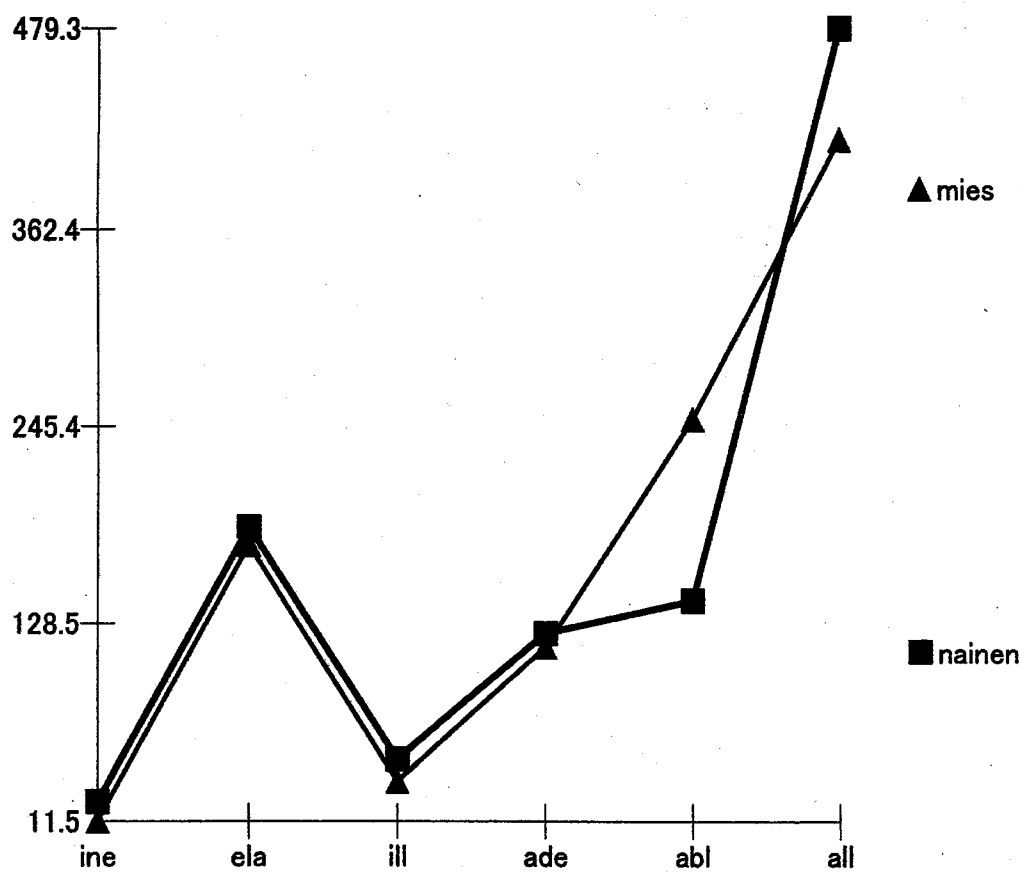


図4 SKTPコーパスにおける mies「男、夫」と nainen「女」の場所格形の出現頻度（理想化+補正）（cf. 表9）

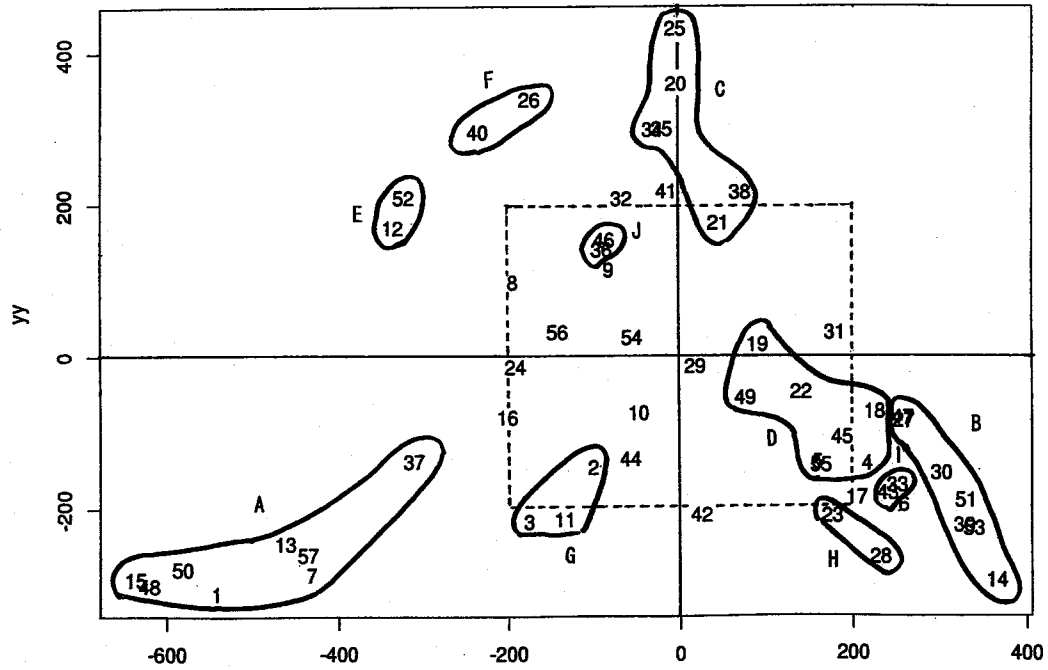


図5：名詞57語の場所格使用パタンの主成分分析

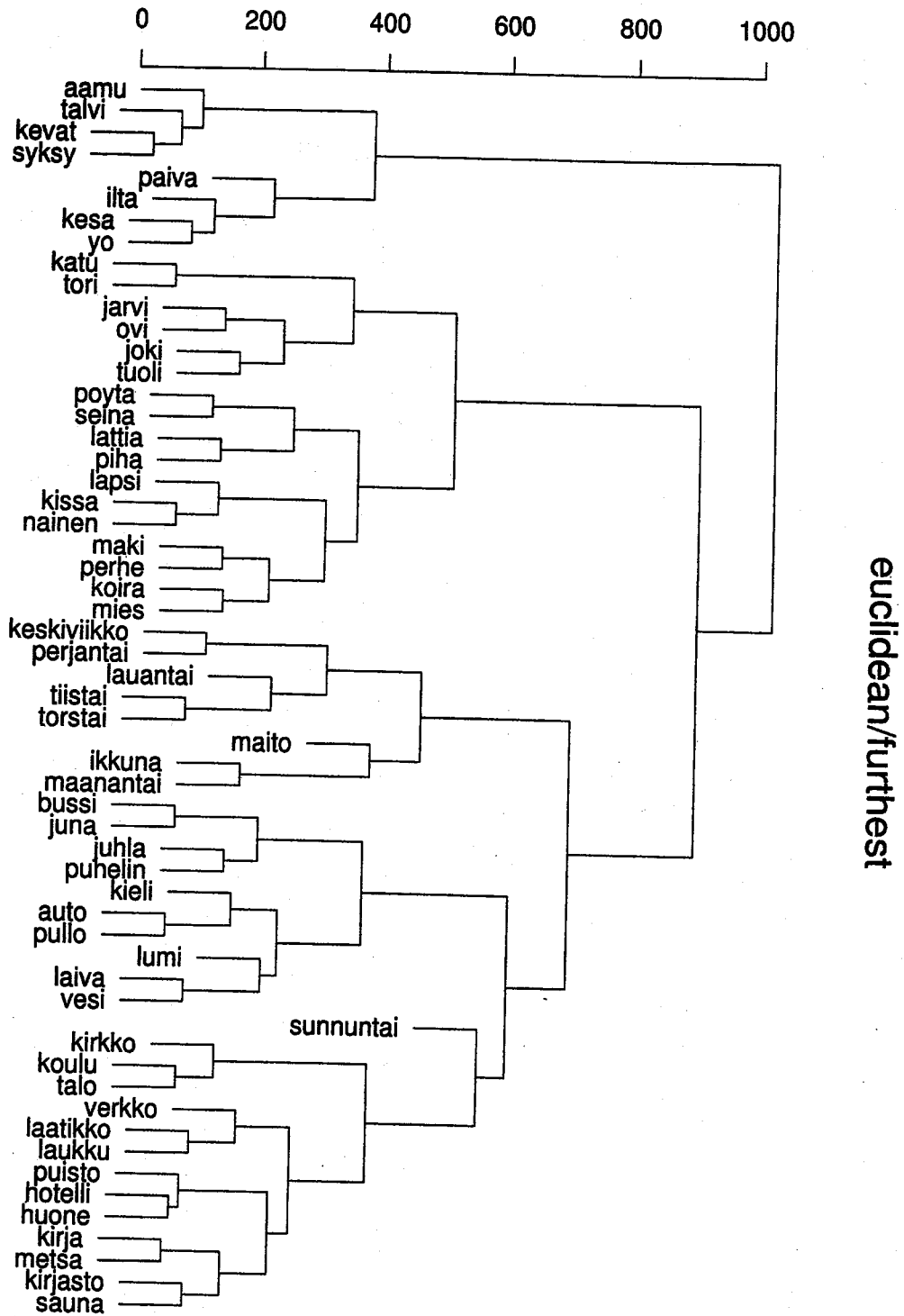


図6：名詞57語の場所格使用パタンの階層クラスタ分析(ユークリッド距離，最長距離法)

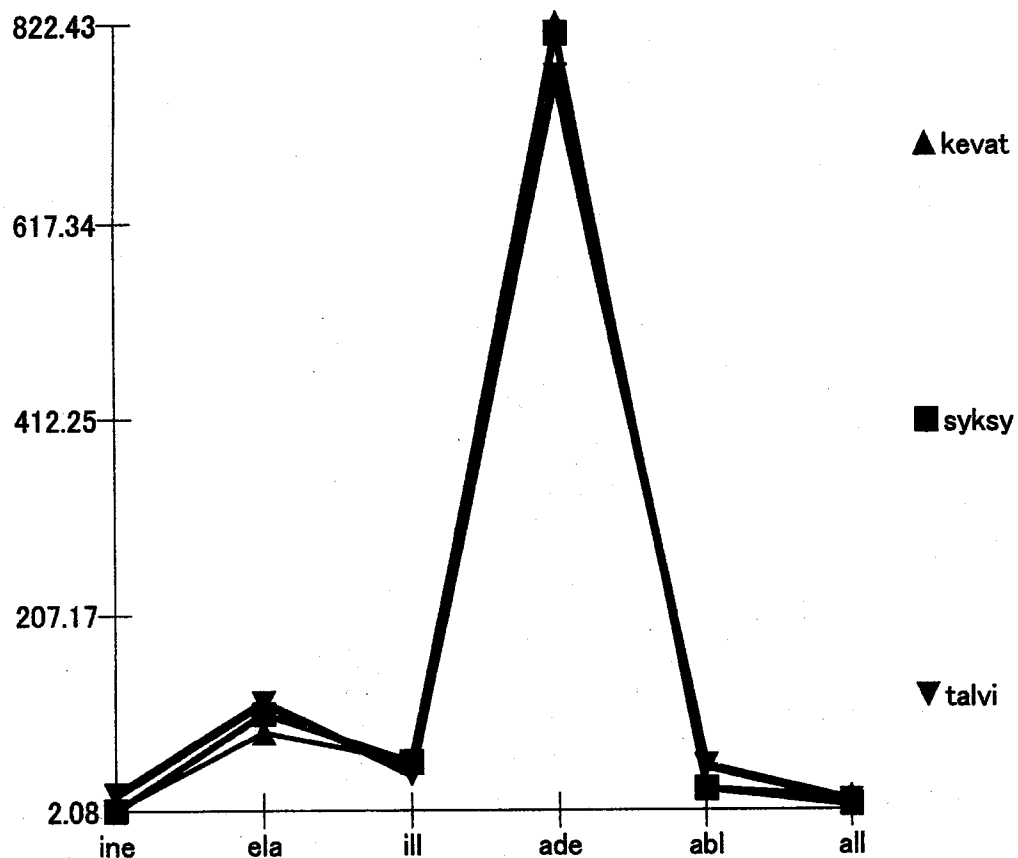


図7 : kevät (15)「春」—syksy (48)「秋」—talvi (50)「冬」

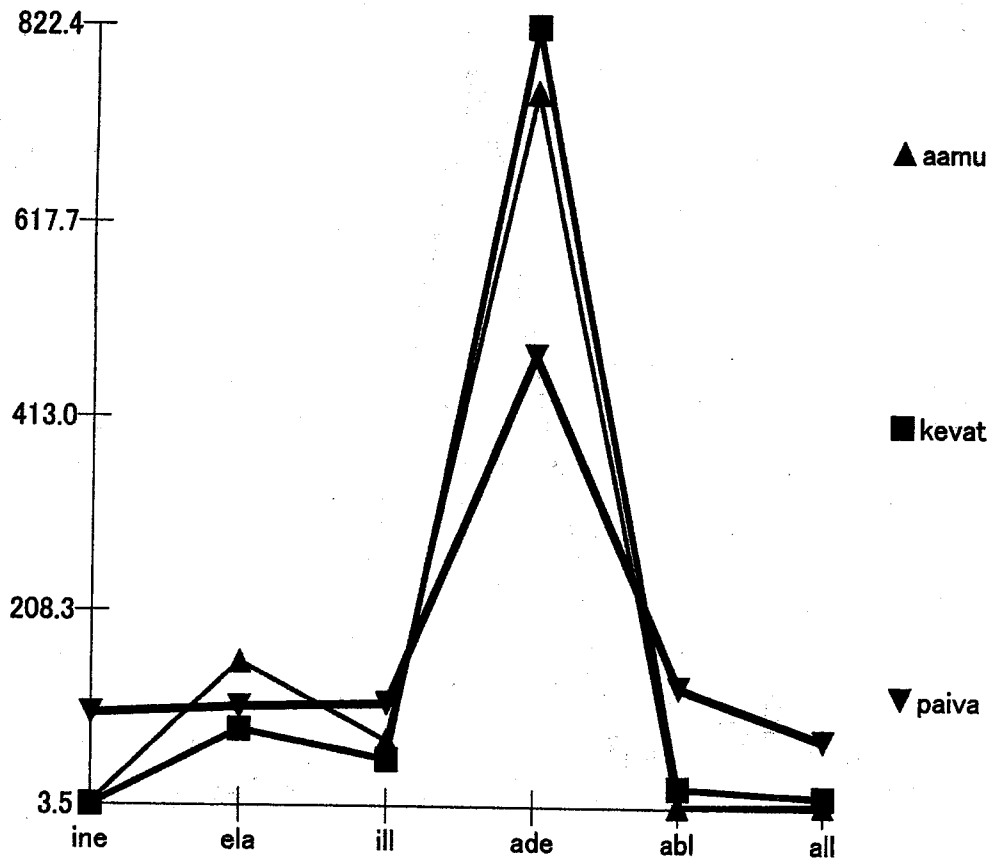


图 8 : aamu (1)「朝」-kevat (15)「春」-paiva (37)「日, 日中」

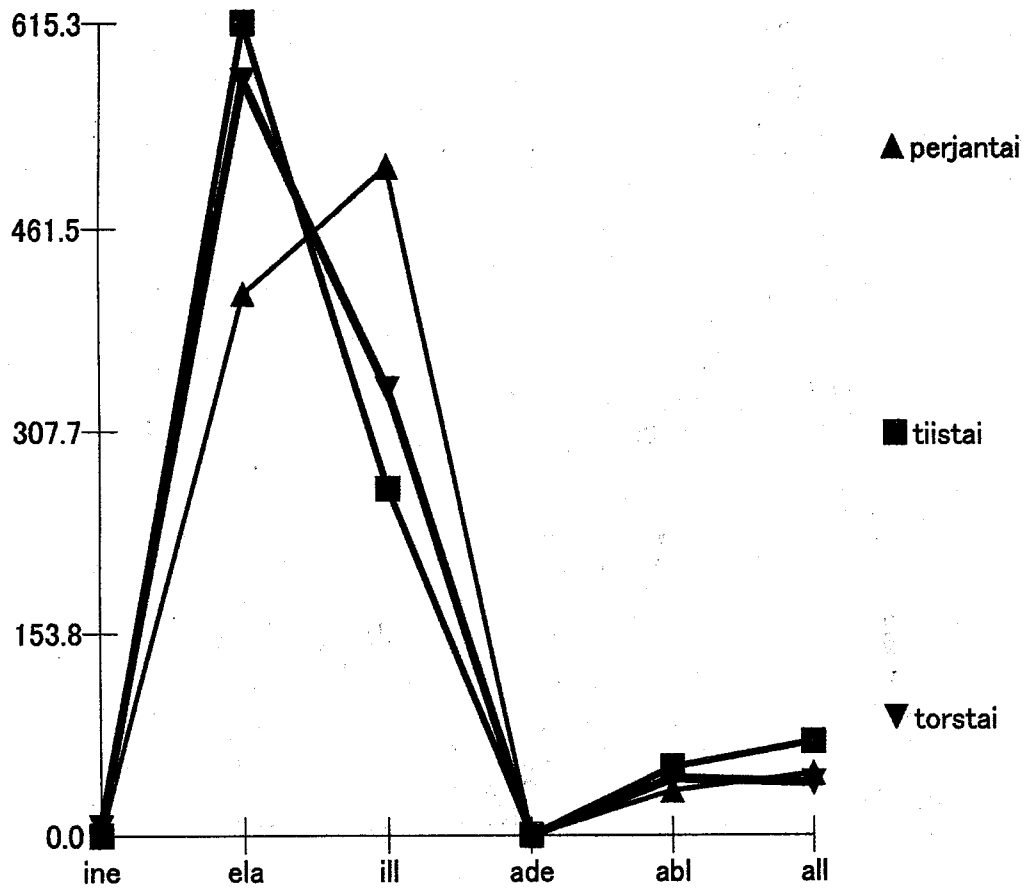


図9 : perjantai (39)「金曜日」—tiistai (51)「火曜日」—
torstai (53)「木曜日」

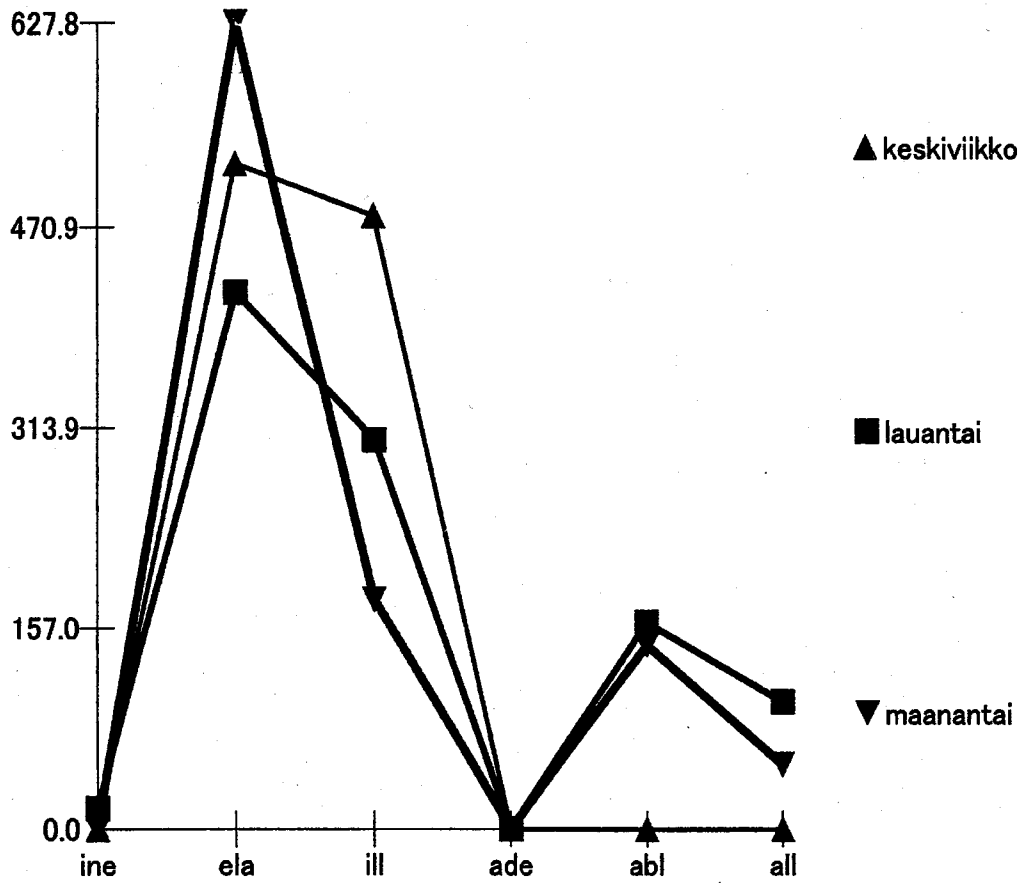


図10 : keskiviikko (14) 「水曜日」 - lauantai (27) 「土曜日」 -
sunnuntai (47) 「日曜日」

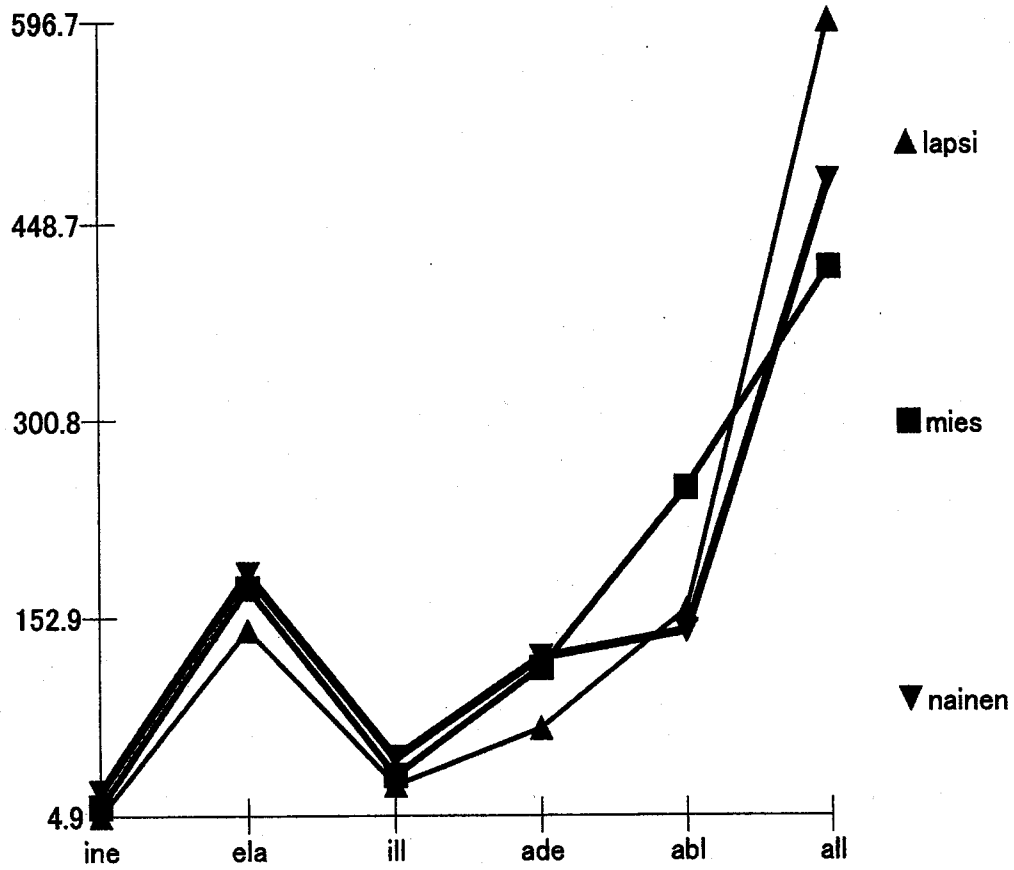


図11 : lapsi (25)「子ども」-mies (34)「男」-nainen (35)「女」

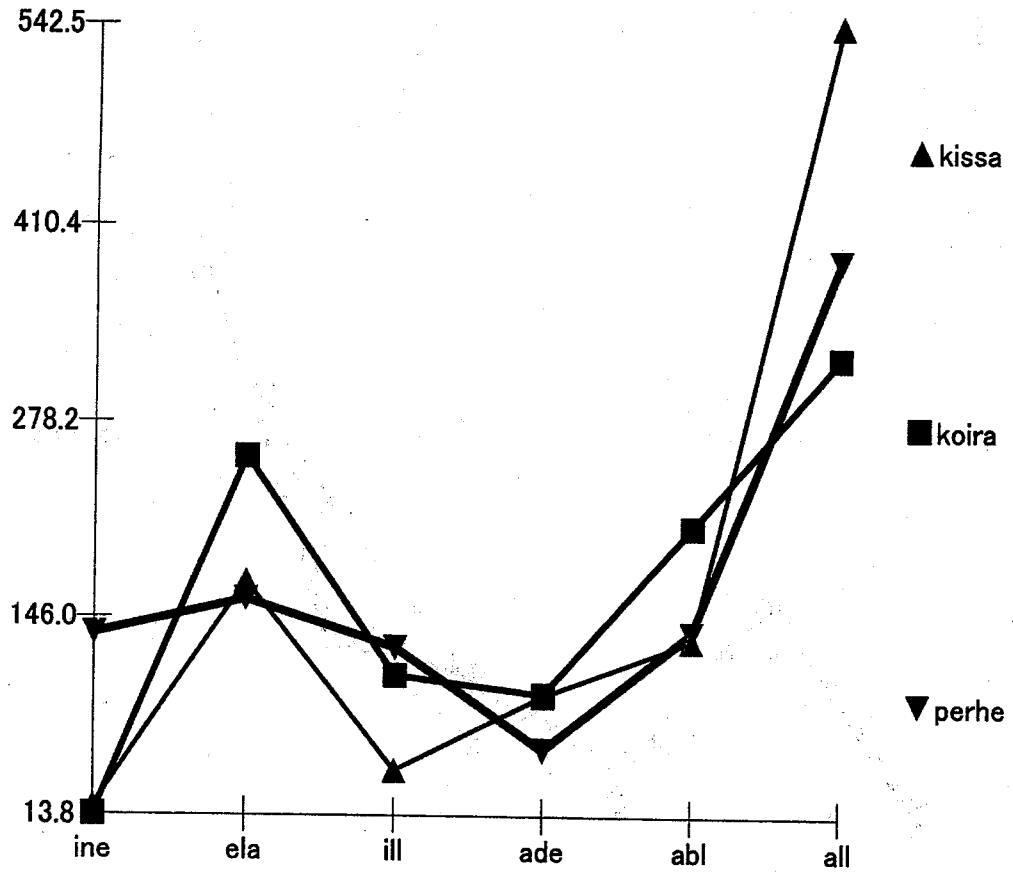


図12 : kissa (20) 「猫」 - koira (21) 「犬」 - perhe (38) 「家族」

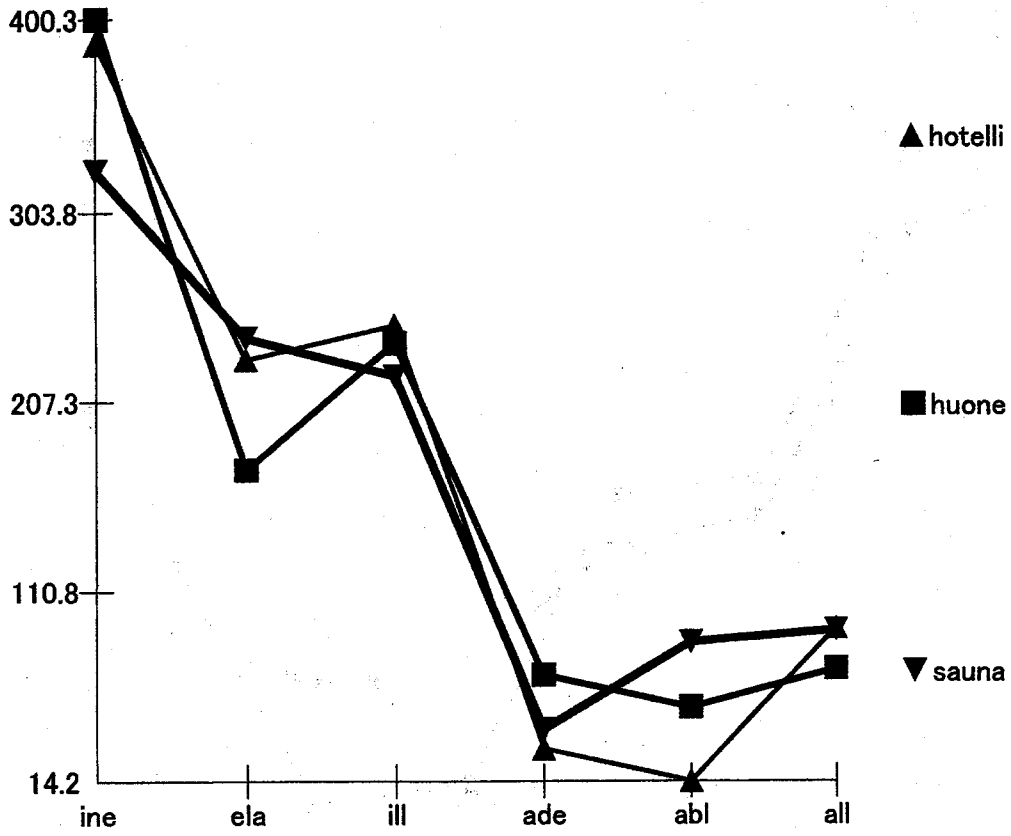


図13 : hotelli (4)「ホテル」 - huone (5)「部屋」 - sauna (45)「サウナ」

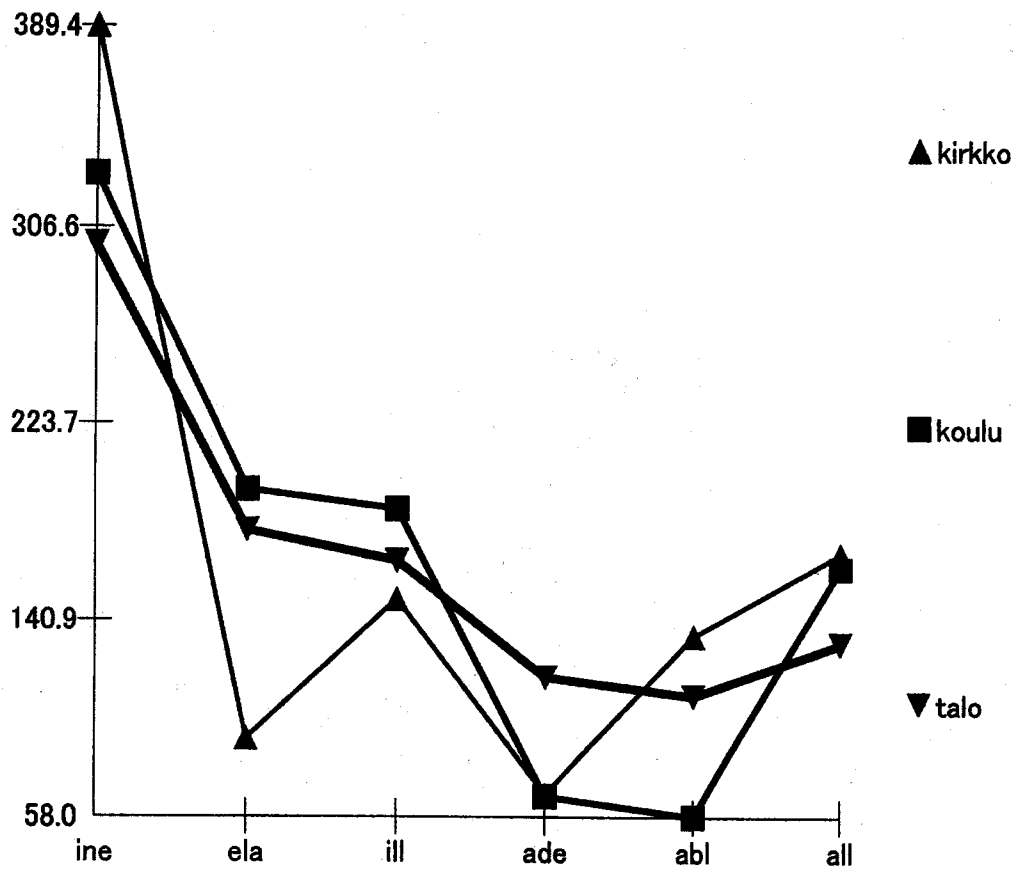


図14 : kirkko (19) 「教会」 - koulu (22) 「学校」 - talo (49) 「家」

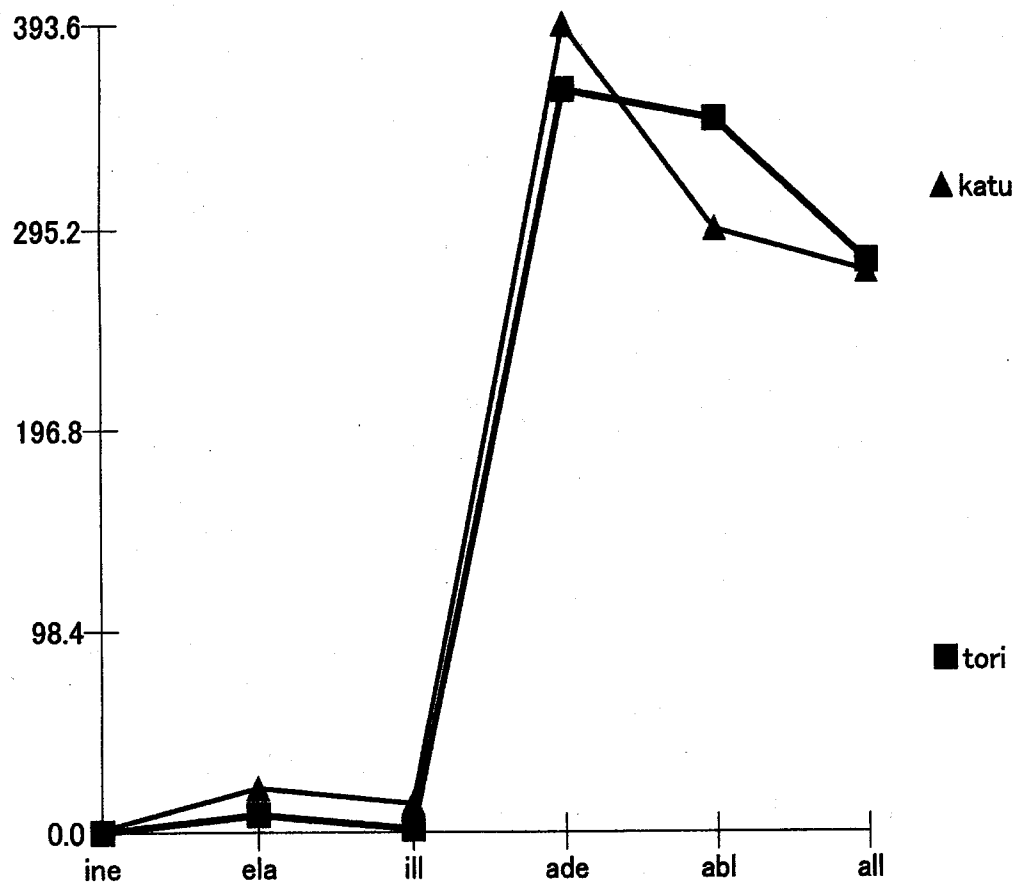


図15 : katu (12)「通り」-tori (52)「広場」

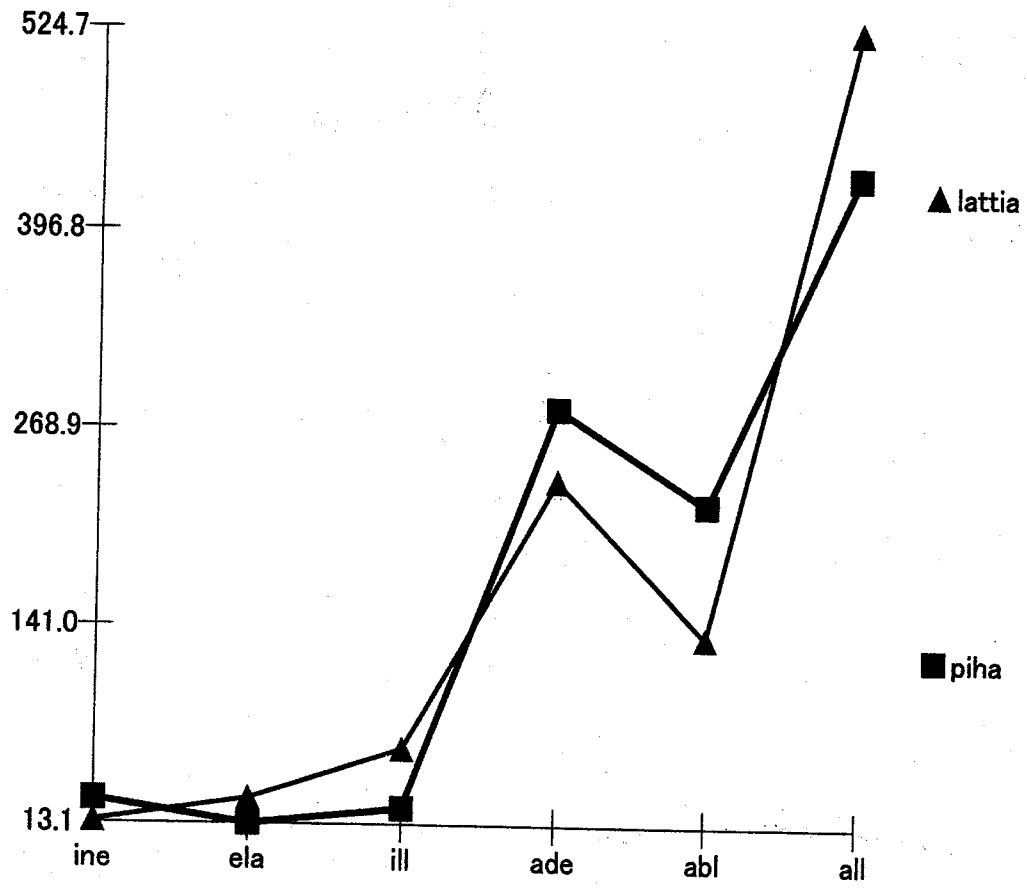


図16 : lattia (26)「床」 - piha (40)「庭」

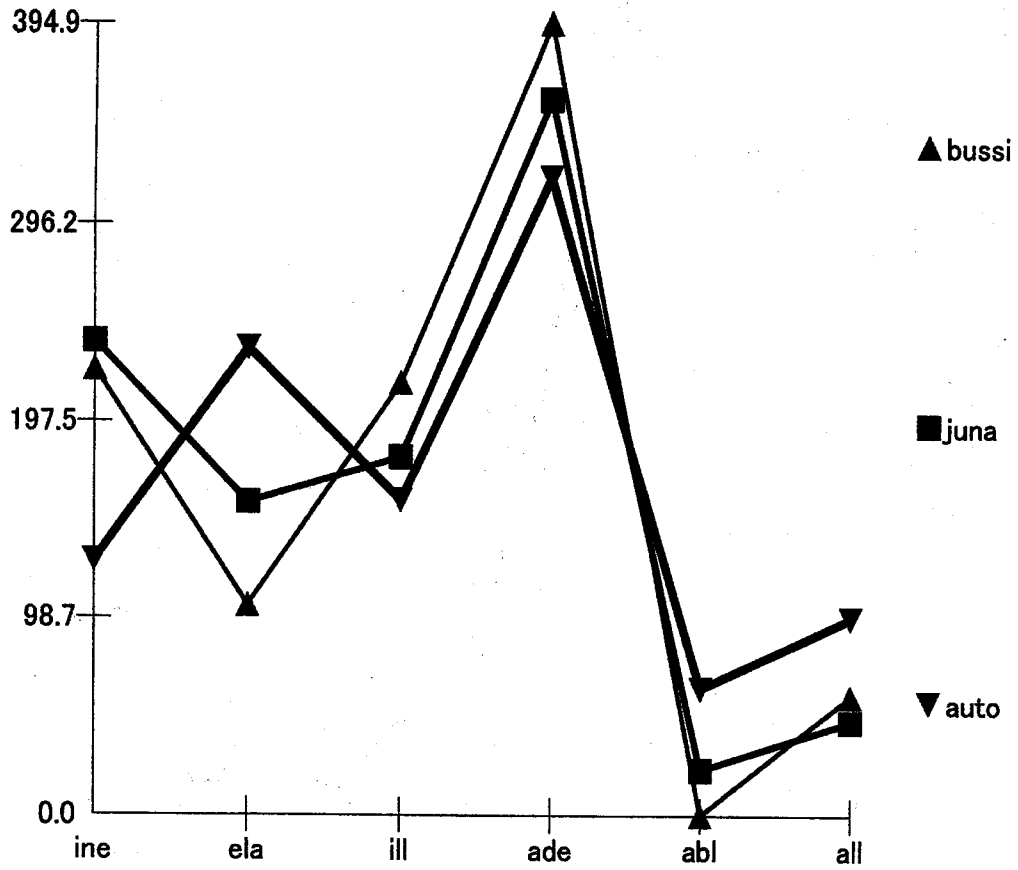


図17 : bussi (3) 「バス」 - juna (11) 「列車」 - auto (2) 「自動車」

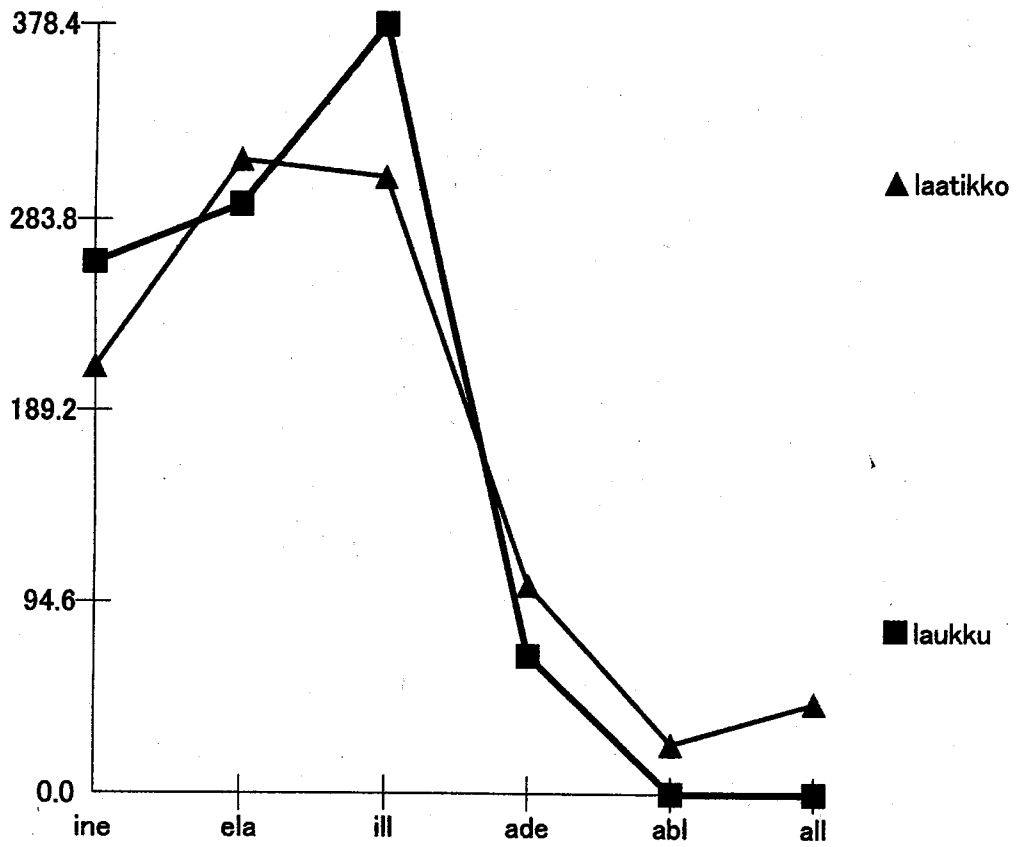


図18: laatikko (23)「箱」-laukku (28)「かばん」

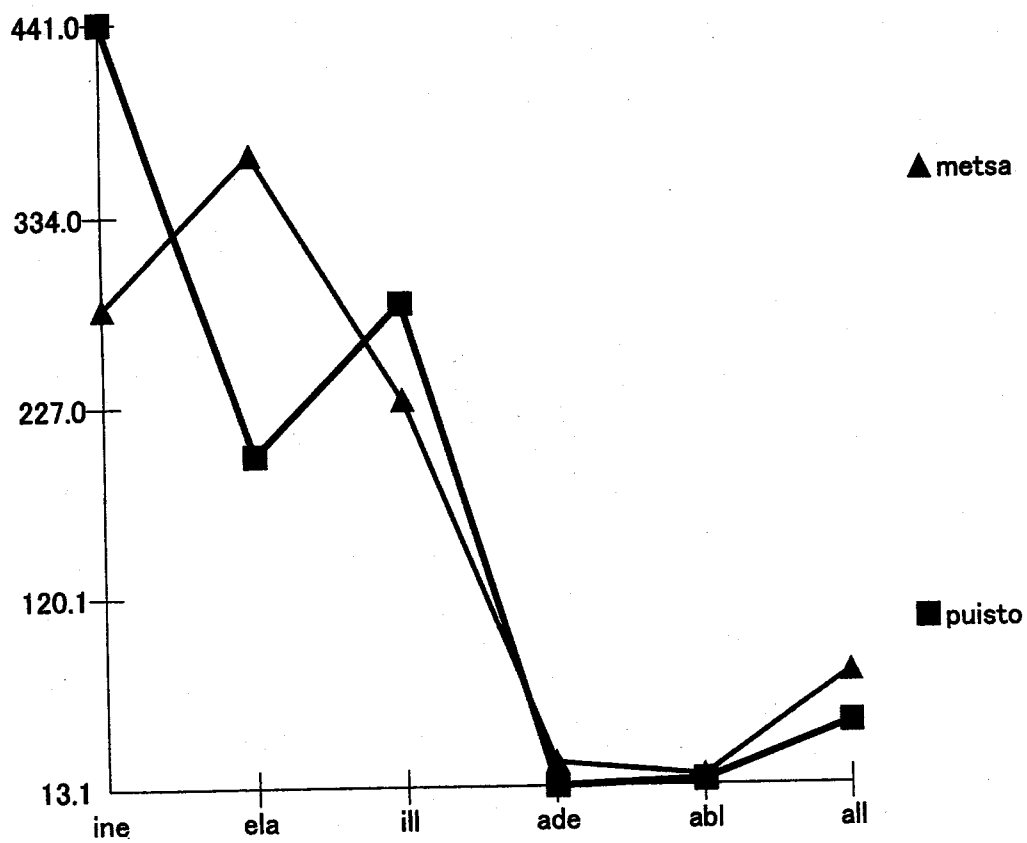


図19 : metsa (33) 「森林」 - puisto (43) 「公園」

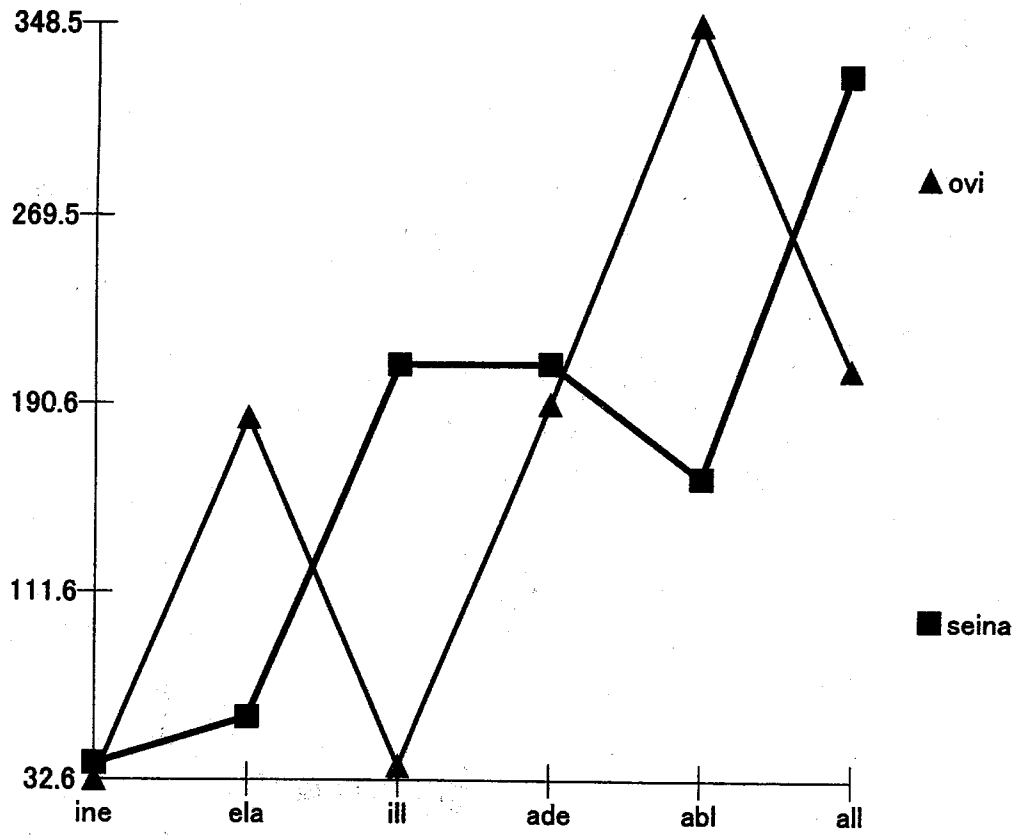


図20 : ovi (36) 「ドア」 - sein (46) 「壁」

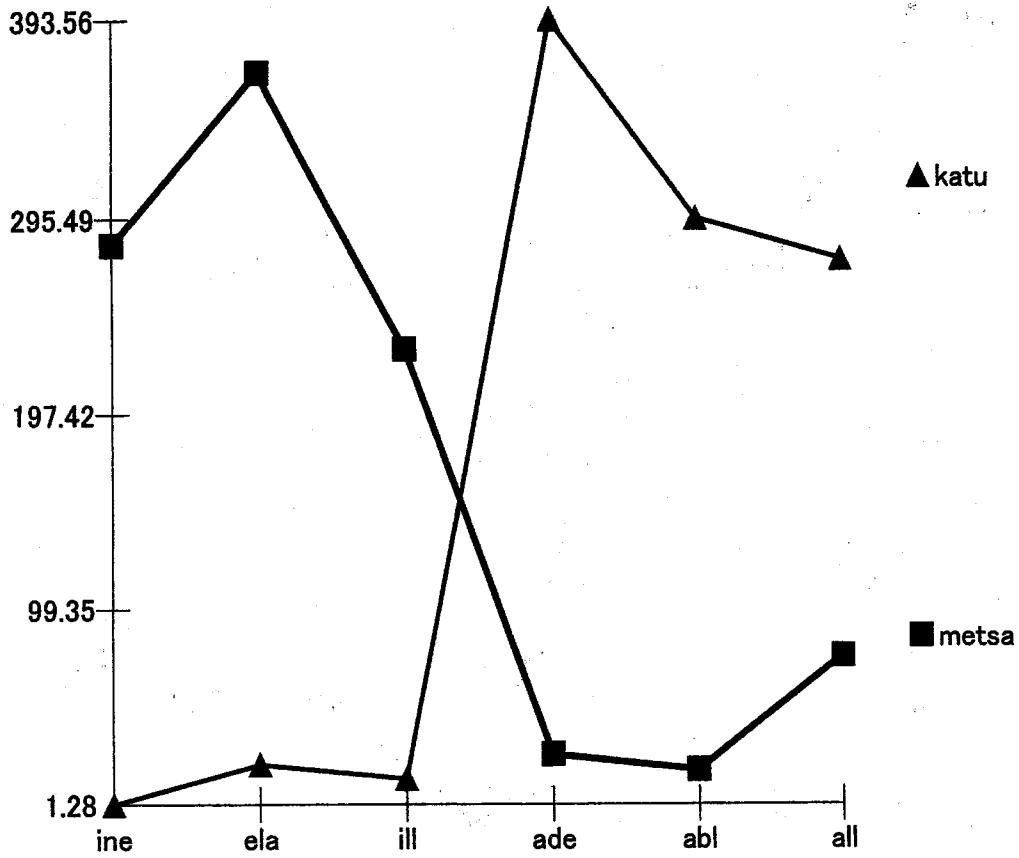


図21 : katu (12) 「通り」 -metsa (33) 「森林」

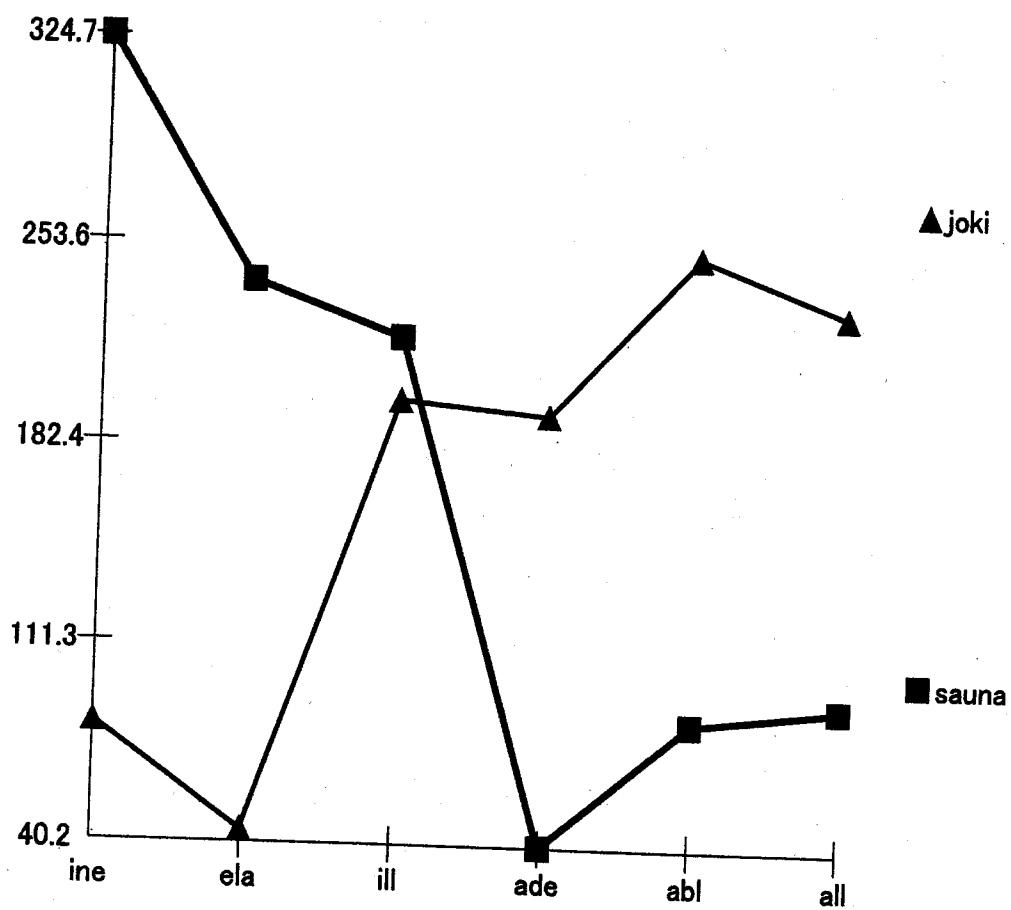


図22 : jarvi (8) 「湖」 - sauna (45) 「サウナ」