DOCUMENT RESUME

ED 041 516 FL 001 856

AUTHOR Mackey, William F.
TITLE Concept Categories as Measures of Culture Distance.
INSTITUTION International Centre for Research on Bilingualism, Quebec (Canada).
PUB DATE 69
NOTE 45p.

EDRS PRICE EDRS Price MF-$0.25 HC-$2.35
DESCRIPTORS Bilingualism, *Concept Formation, Contrastive Linguistics, *Cross Cultural Studies, Cultural Background, Cultural Differences, *Cultural Interrelationships, Cultural Traits, *French, *Language, Modern Languages, Nominals, Sociocultural Patterns, Sociolinguistics

ABSTRACT

The research reported in this paper tests the possibility of using differences in conceptual categories in the measurement of differences between two cultures. In the initial experiment, which contrasts French and French-Acadian linguistic habits, the authors limit themselves to conceptual categories reflected in concrete nouns representing the naming or labeling of things, groups of things, and observable actions. An overview of the contents includes: (1) delimiting the problem, (2) coding conceptual categories, (3) variations in time and space, (4) indices of diversity, (5) indices of intensity, (6) selecting conceptual fields, (7) sampling the population, (8) recording measurable responses, (9) analyzing and computing the data, and (10) appraising the results. Some 3,000 persons contributed more than four million word-token responses for analysis. Tables and graphs illustrate findings. (Author/RL)

centre international de recherches sur le bilinguisme

CONCEPT CATEGORIES AS MEASURES OF CULTURE DISTANCE

by WILLIAM F. MACKEY

1969

CIRB
ICRB

Documentation internationale et recherche interdisciplinaire dans le domaine du bilinguisme, du biculturalisme et des phénomènes connexes

International documentation and interdisciplinary research in the field of bilingualism, biculturalism and related phenomena

Centre international de recherches sur le bilinguisme
International Centre for Research on Bilingualism
Université Laval, Cité universitaire,
Ste-Foy 10e (Québec) Canada

CONCEPT CATEGORIES AS MEASURES OF CULTURE DISTANCE

William F. Mackey

International Centre for Research on Bilingualism

Outline: *Overview. 1. Delimiting the problem. 2. Coding conceptual categories. 3. Variations in time and space. 4. Indices of diversity. 5. Indices of intensity. 6. Selecting conceptual fields. 7. Sampling the population. 8. Recording measurable responses. 9. Analysing and computing the data. 10. Appraising the results.*

Overview

The purpose of the research described here is to test the possibility of using differences in conceptual categories in the measurement of differences between cultures.

After a short delimitation of the problem and a definition of the role of conceptual categories in culture and their variation in time and space, we shall explain the choice of categories and cultural areas used in the study, the tests used to elicit the categories, field work, accumulation and computation of data. Finally we shall explain the methods used to isolate conceptual categories and measure their differences in intensity and diversity as possible indices of differences between cultures.

1. DELIMITING THE PROBLEM

All creatures, including man, must live in some sort of environment. In order to exist in this environment they must devise ways of dealing with it, since they may both be affected by it and have an effect upon it. The environment of man is varied and changeable; to deal with it he has evolved a traditional guide to behavior, which includes what he eats and wears, how he gets his food and clothing, the tools and techniques he uses, his

relationships with his fellows and the institutions, customs, laws and beliefs which enable him to work with his fellow man. This guide to behavior, enabling man to deal with his environment, and all it includes, is largely acquired and retained through traditional conceptual categories most of which may be transmitted through language. The language provides the acoustic or visual forms through which categories are coded for use in communication. It is through the language that we can get at a conceptual category, which is simply one of many possible ways of arbitrarily grouping objects, events and other phenomena so that one can think and talk about them.

Most languages, for example, have conceptual categories for the three forms (solid, liquid and gas) of life's most basic substance -- water. But not all have the same type and number of categories. For the solid form, most European languages distinguish only between ice and snow. Languages traditionally used in arctic environments, however, distinguish more types in their code. In Baffin Land, in 1968, this writer was able to identify 21 distinctly different words for types of snow and ice, some words requiring long sentences to translate into a European language. Snow and ice phenomena represent richer conceptual

categories in Baffin Eskimo than they do in English, and these are more highly coded. It would seem, therefore, that conceptual categories, as manifested in the code of the language, might be used as an indication of the importance which the phenomena they represent may have in the environment of the people traditionally using that language.

2. CODING CONCEPTUAL CATEGORIES

In coding conceptual categories a language has a number of devices at its disposal -- words, grammatical units (prefixes, suffixes and infixes, function words), word-patterns and even intonation features. But the number of conceptual categories which can be coded by phonetic or grammatical units (forming a closed system) is limited because there are proportionally far fewer such units in any language than there are units of vocabulary which, forming an open system, are also more adaptable to changes in environment. It is therefore the vocabulary rather than the pronunciation or grammar of a language which is the most usable indication of the conceptual categories coded by the language.

These conceptual categories coded in the vocabulary, in the nouns, verbs, adjectives, adverbs and other form-classes,

cover everything that man can see, feel and imagine. Many of the words in a language are labels for categories of notions, ideas, attitudes, beliefs, relationships and judgments which have no direct counterpart in the physical world. Although the categories of non-physical phenomena may be as indicative as anything else of the differences between various cultures (or guides to mental and physical behavior), they are hardly the best type of conceptual categories to be used initially as a test of the validity of such categories as a measure of differences between cultures. This is because, not being observable directly, their relationship with the unseen would be difficult to prove. In order to prove a measurable relationship one needs to deal with the most observable items of a culture -- the objects used in the environment. These are usually expressed in concrete nouns -- or equivalent form-classes.

We have therefore limited ourselves, for the purposes of this initial experiment, to conceptual categories reflected in concrete nouns, representing the naming or labeling of things, groups of things and observable actions.

3. VARIATIONS IN TIME AND SPACE

These conceptual categories are by no means permanent; they vary in time and space. Old words disappear and new ones are born. Old words become labels for new, expanded or different concepts. Conceptual categories may take on new shapes and sizes as the environment which they reflect evolves. The most unstable areas of human environment represent the most unstable categories; words for parts of the body are more stable than words for things used to clothe them. Conceptual categories used in one area would presumably change if the people using them were to settle in another. The change would be a measure of the difference in the environment. To test this hypothesis it is useful not only to limit oneself to the use of categories of the concrete, that is, to the names of things, but also to make sure that the names and categories were the same in one point of time and space, and have become different in another point of time and space. It would therefore be preferable to begin with what might be loosely described as the adaptation of a single language to the evolution of two cultures.

Several colonial and immigrant languages could be used as a test. We have had the opportunity to select French as it is used in Europe and North America. The part of our data used for this

first experiment comes from France on the one hand and Acadia on the other.

The Acadians, who came originally from France, settled in North America in the first part of the 18th century, especially in the area known today as the Canadian Atlantic Provinces. When the first settlers to Acadia left France, it may be assumed that they had the same available names as did their compatriots who stayed behind (e.g. *vin* may have been high on the list of beverages and *carosse* on the list of vehicles). As they became accustomed to their new environment, their culture changed in a direction different from that of their compatriots, resulting in new names or new categories of things with new importance. New words came in; others disappeared. Old words took on new meanings and functions. Since wine, for example, became a little available drink (no grapes), the word *vin* became less important. The compatriots back home also changed with the passage of time; but because they were in a different environment their culture evolved in a different direction. They maintained some of the forms whose importance had diminished in America, but they also rejected some of the same categories.

Thus, after some 200 years of complete separation and absolute lack of contact, the two groups no longer belonged to the same and identical culture. This difference was accentuated by the total immergence of the Acadian group into the North American English-speaking community, first by dispersion and second by the necessity of co-existence. Yet both the Acadians and the descendants of their ancestors in France continued to speak the same mutually comprehensible language -- two dialects of Modern French.

A study of the conceptual categories of these two groups should help us find out whether any differences in conceptual categories resulting from the transplantation of a language in a new environment are measurable and, if so, what variables may be measured and by what means. If it can be established that conceptual categories may be used as measures of differences between cultures conveyed by two dialects of the same language, it would be reasonable to assume that the same would be possible for two cultures using two different languages.

These differences, it is assumed, will be reflected in the names used to segment the realities of the environment and the words used most readily -- those which come to the minds

of most people in the community. These words and the people using them can both be counted.

The names or segments of a conceptual category in one culture can be computed in comparison with those of another culture, giving an index to the diversity of names which each possesses. The number of people bringing these to mind may be counted, giving an index to the intensity with which a category is represented in each culture. Diversity and intensity, therefore, are the two indices we will attempt to elaborate for the measurement of intercultural distance.

4. INDICES OF DIVERSITY

Diversity is a measure of the number of conceptual categories in which one culture divides its inner and outer environment in comparison with another culture. This can be measured by counting the number and extent of such distinctions. It may be assumed that the higher the number, the greater the distance between the cultures in those areas in which such distinctions occur. Differences may also appear in the number of labels a culture gives the same thing. Finally, in cases of culture contact, a conceptual category may be modified by association with another culture, may disappear, or be replaced.

In measuring the degree of diversity in a category present in one culture in comparison with that of the same category in another culture, it is necessary to establish an equivalence between what may be labeled in one culture with one name and represented likewise or differently in another culture. This equivalence may appear in seven different forms.

i) Two cultures may have the same name for the same thing; *chapeau* refers to *a hat* in both France and Acadia $(X = X)$. ii) The two cultures may have the same name (X) for two different things (x,a); *slip* is male underwear in France, female in Acadia $(X^x = X^a)$. iii) The two cultures may have different words for the same thing; *crêpe de sarrasin* in France may be *plogue* in Acadia $(X = A)$. iv) The two cultures may have different things and different names; there is no French equivalent for the Acadian *poudrerie* (sweeping snow storm) or for *sloche* (slush) $(X = O)$. v) A word representing one conceptual category in one culture may be equivalent to several conceptual categories in another. French *lueur* covers English *glow, gleam, glimmer, shine* $(X = \frac{A}{B})$.

vi) The same word may cover a larger semantic category

in one culture than in another. The word *administration* covers more ground in French than it does in English. The word *canadienne* refers only to persons in Acadia; in France, it is also the name for a sort of car-coat ($X^x = \frac{X^x}{X^a}$). vii) One word referring to one thing in one culture may be equivalent to several words referring to the same thing in another ($X = A, B$).

5. *INDICES OF INTENSITY*

Since the continuum of culture is not co-terminous with either language or nationality, it is not sufficient to say that a conceptual category exists in one culture and not in another -- either with one or many names. One must also know how important the conceptual category is in a given culture at a given time. This importance can be measured by counting the number of people for whom the concept is available, measuring the readiness with which this availability is made known. This use of availability as a measure of the intensity of conceptual categories within a culture is based on the seven following postulates.

i) People are more familiar with objects that form an important part of their culture than with those that do not.

ii) People tend to name familiar categories more readily than unfamiliar ones. This is easily demonstrated by asking a hundred people to name any part of the body; more will name *nose* than *pancreas*. The more important an item in a culture, the more it is reflected in the availability of the labels. The more important the item of vocabulary, the more readily it comes to mind.

iii) The readiness with which persons name objects (availability) is a measure of their cultural importance to such persons. And differences in the degree of cultural importance are reflected by differences in degree of availability.

iv) The greater the difference between cultures in the availability of a label, the greater the difference in the role of the object which the label represents. If the word *wine* is 100% available in Culture A, and 25% in Culture B, it may be assumed that the beverage "wine" is in fact proportionally more important in the lives of the people of Culture A than it is in that of those of Culture B. But it could be more important

in B than in C, which may give it only 1%. The greater the difference, the greater the cultural distance.

v) In any two cultures, the number of objects will differ in degree of availability. The more objects that differ, the greater the distances between cultures. Intercultural distance is a function of the number of different objects and the differences in availability of mutually known objects.

vi) In each conceptual field the more subjects listing a word, the more important the word within that field. The more words in a conceptual field, the more important the field.

vii) Differences in the degree of availability may be used to measure differences in cultural importance.

Availability of categories cannot only be used as a measure of distance between cultures, but also distances within the same culture over a period of time.

Since intercultural distance is also a function of object-grouping (categories) such categories must be made equivalent

before their degree of availability can be compared. Equivalent categories must first be established for each item tested. *(See above.)*

What is being measured, therefore, is the degree of availability of conceptual categories in two cultures, using dialects of the same language, in an attempt to discover the extent to which these figures may be used as indices of the difference between two cultures.

Degree of availability is a measure of the intensity of interest which a conceptual category has for a culture. It can be measured by counting the number of persons ready to include a category within a given field.

But what are these fields and what persons are to be counted? What categories must be isolated, how should they be obtained, how many and from whom? These questions may be treated under three main headings: the selection of the conceptual fields, the recording of measurable responses and the analysis and computation of data.

6. SELECTING THE CONCEPTUAL FIELDS

If we were to take all the things in the world -- chairs,

trees, houses, etc, and arrange them in the order of their commonality, we could have a scale that might include such things as the parts of the body at one end and certain local specialized tools at the other. For there are certain things that all men everywhere are likely to have in common and other things that are limited to a few men in a few places. The scale would correlate with degrees of differences in availability. In other words, there would be some things where the similarity in degrees of availability would be high, reflecting cultural commonality (e.g. parts of the body) and other things where it would be low (e.g. certain tools), reflecting cultural divergence.

For the purposes of this study, all these things are grouped into a couple of dozen conceptual fields, into which the most common concepts may be the most conveniently classified. For the common physical objects, these conceptual fields include food, clothing, housing, transport; for common actions they include such fields as work and play. It is within a given conceptual field that the availability of a concept may be measured. And since the measurement has to be based on the role of the concepts within a community, the next problem is to determine who should represent the culture in the community.

7. SAMPLING THE POPULATION

The most obvious population sample would have been a structured one -- a cross-section of the entire adult population. But because of doubts about the representativeness of adults and uncertainty of the sampling; and also because of the difficulty in obtaining large enough samples, school children between the ages of nine and eighteen were considered preferable for the purposes of this initial experiment.

Children supplied the general concepts needed for the study, without the added complication of technical vocabularies which the adult population of a technologically developed country is likely to supply. Words for parts of the body supplied by nurses, doctors and medical students are likely to be more specific than those supplied by cooks, accountants and law students. Choice of adults would have obliged us to structure the sample -- a task not deemed to be necessary in a preliminary feasibility study such as this one.

A total of some three thousand persons between the ages of nine and eighteen were therefore used as a sample of the populations.

From this population it was necessary to elicit responses within the selected conceptual fields in order to obtain measures of the intensity and diversity of the conceptual categories. How was this done?

8. RECORDING MEASURABLE RESPONSES

The readiness with which a person names an object is a measure of its importance to him; the more persons who name a given object in a community, the more important that object is likely to be for them.

One way of determining which concepts were most familiar would have been to record a sufficient number of speech samples and count the words which appeared most often. As a matter of fact, our experiment began with a field recording in Acadia of some hundred hours of unrehearsed conversation. This turned out to be of little use, because of the instability of the frequencies of the most common names like *tête, chapeau, manteau, nez.* Persons often used these names only when an unusual rather than usual event occurred.

It was noted for example that people, who used their tongue every day of their lives talked about this part of their body only if something unusual happened to it -- when for instance they happened to burn or bite their tongue. Frequency of occurrence in daily usage, therefore, turned out to be a poor measure of the availability of concrete objects. Moreover, frequency of occurrence in daily usage had to be based on the analysis of discourse -- of the messages produced; we were after the code that could produce the messages.

We therefore used a type of controlled, conscious and structured word-association device, asking people to give names for objects belonging to a certain conceptual field. This decision also enabled us to replicate availability studies done elsewhere and to make use of some of the results. We had our subjects from each region supply, under controlled experimental conditions, the names of objects and actions that came within each conceptual field. They were given a quarter of an hour per field, each person devoting a maximum of ten hours to a dozen fields. This yielded several million word-token responses.

9. ANALYZING AND COMPUTING THE DATA

A total of some three thousand persons supplied more than four million word-token responses. Of this mass of data only that

supplied by some 1,600 subjects (aged 9 to 13) in France and Acadia has been analysed and computed, and that for only sixteen conceptual fields. It is from this computer analysed data that the conceptual categories studied here have been isolated, grouped, analysed, measured and compared. Results for four conceptual fields are presented here, namely, clothing, food, transportation and pastimes.

The conceptual analysis and computation was done in three main phases:

Phase 1: Computer treatment of raw data. More than a million word-token responses were integrated into some 40,000 word-types. These were arranged in decreasing order of availability according to conceptual field.

For each conceptual field, a list (Table 1) was provided indicating for each word-type: a) its rank-order (Column 1), based on b) the total combined population out of 1,600 (700 Acadian and 900 French subjects) supplying the word (Column 2), c) the percentage of each group supplying it in Acadia (Column 4) and in France (Column 5), d) the superiority of Acadian over

French (A - F) in the percentage of subjects supplying the item (Column 6); or the superiority of French over Acadian, also percentage-wise (Column 7). For the results for each of the four fields, see Table 1 (1.1 Clothing; 1.2 Food; 1.3 Transportation, and 1.4 Pastimes).

Phase 2: Isolation and grouping of conceptual categories. Word-types for conceptual categories in each field were isolated and grouped. Category equivalences were established so that similarities and differences could be measured. Results were computed for four levels of categorization: 1. Conceptual Field (e.g. clothing); 2. Conceptual Category (e.g. head-dress); 3. Category Segment (e.g. hood); 4. Names (e.g. *capuchon, capuche, chaperon, capeline*). An example of this is given in Table 2. Note how differences tend to become more pronounced as categorization becomes more specific. Over-all difference in the field of clothing is less than 1%; whereas in some of the names of hats, it exceeds 40% (e.g. *chapeau*). The rank-order indicates the position of the word on a list of combined frequencies (Acadian + French)

in decreasing numerical order. So that *chapeau* which is 8th in the order of importance in the field of clothing, has a combined frequency of 921 votes out of a possible 1,600 (700 Acadians + 900 French subjects) (See Table 1.1).

Phase 3: From the results of the first and second phases, differences in diversity and intensity were calculated.

Diversity was calculated by counting the number of different words in each conceptual category. From Table 2, for example, we count eleven words used in Acadia for head-dress and twelve in France. Thus the index of diversity for head-dress is Acadia: 11; France: 12. We might perhaps conclude from this that the variety of head-dress in France is slightly higher than it is in Acadia. Whether this indicates that head-dress as such is used more in France than it is in Acadia is not certain. The converse may be true if the intensity indices are an indication. These, based as they are on availability figures, referring again to Table 2, give 53.4% for Acadia and 46.6% for France. The difference both ways may not be

significant, since the correlation between intensity and diversity, if one does exist, has not yet been determined. In the meantime, it is worth noting that the diversity of labels and distinctions in one culture may be the result of culture contact, as is here exemplified in the number of loan words and neologisms appearing in Acadia, compared with the number appearing in France (See Table 3.) It is the intensity index, however, which seems to be the more reliable indicator of cultural difference, if the selection shown in Table 4 is typical.

These four tables, however, are not structured enough to permit one to grasp the extent of the difference between the two cultures. What is needed is some sort of statistical model capable of absorbing and visually organizing the results into easily observable patterns. This is what we have attempted to do in Figure 1.

The important thing in Figure 1 is the over-all pattern or spread revealed by the integration of the differences. We note that the pattern of the spread differs with the conceptual field. In using this type of representation, identical cultures would have a difference of zero, revealing a perpendicular line; entirely

different cultures would have a consistent difference of 100%, appearing as a horizontally lined rectangle. Between these two extremes are triangles of different shapes and sizes, indicating the amount of difference between cultures for each conceptual field. When all conceptual fields will have been measured thus, it would be a simple matter to integrate all differences into a single pattern, enabling us to see at a glance the over-all distance separating two different cultures.

On the model of Figure 1, we plotted the results indicated in the last two columns of Table 1, showing the difference in availability for the first hundred words, arranged according to conceptual field. The results listed in Column 6 (Acadia - France) were plotted on the left-hand scale; those in Column 7 (France - Acadia) on the inter-facing right-hand scale, both in rank-order. This gave us, for each conceptual field, a pattern of difference based on the hundred numerically most important labels (word-types), each representing the responses of not less than 95% of the population. The remaining word-types, ranging in number from 338 labels for types of clothing to 1,228 for pastimes, were each supplied by less than 5% of the population of the respective areas, that is, by less than 35 subjects out of a possible total of 700 in Acadia and by less than 45 out of a possible 900 in France.

The patterns in Figure 2 were based on statistics taken directly from availability indices of word-types. What would the difference in pattern be if we were to base the variables on the derived conceptual categories ? This is what we have attempted in Figure 3. Here our sources were the categories exemplified in Table 2. In Figure 3.1, for instance, we find that the conceptual category of head-dress rates highest in the similarities that exist between Acadia and France (i.e., the least Acadian dominance); the difference, as can be noted by the length of the line, is less than 9%. This has been calculated from the data supplied in Table 2.2 (Conceptual Field: Head-dress -- Acadia, 53.4%; France, 46.6%): 53.4 - 46.6 = 6.8%, which is the amount plotted for head-dress on Figure 3.1. Comparing Figures 3 and 2, one notices such a similarity that one wonders whether there are cases where cultural distance might not be measured directly from availability indices.

After having examined some of the results, we are now in a position to ask whether such a method of investigation and analysis could be used to measure differences between cultures.

The fact that many obvious differences, directly verifiable in practice, are borne out by corresponding differences in the

figures is some indication that the results may be reliable. If we take a look at Table 4, for example, we note that conceptual categories like wine, beret, smock, metro, and soccer, which one would expect to be dominant in France, are dominant in the percentage figures appearing in the table; whereas those for hockey, sleigh, skate, juice, and bingo, expected as dominant in North America, are likewise borne out by the Acadian figures. If the differences are correctly measured in these conceptual categories, why should they not be equally valid for the others? Such results seem to indicate that it might be useful to continue the development and refinement of such measures.

Before such measures can be used with efficiency, however, a great deal of validation, standardization and optimalization remains to be done. Some of the questions which still remain to be answered are: i) How many conceptual fields are needed and what are they? ii) What is the size and make-up of the optimal population sample? iii) What is the optimal number of responses for each conceptual field? iv) What is the optimal age-range? v) What is the most efficient and reliable method of establishing category equivalence?

These are some of the main questions which remain to be answered before this method of intercultural measurement, if valid, could be offered for general use. The encouraging results so far obtained would indicate that answers to these questions may be worth seeking.

TABLE 1

1.1 CLOTHING
(The first hundred out of 438 word-types)

Col. 1 Rank	Col. 2 Population	Col. 3 Name (word-type)	Col. 4 Acadia	Col. 5 France	Col. 6 A-F = %	Col. 7 F-A = %
1	1594	chemise	100.0	97.4	2.6	
2	1336	manteau	94.7	77.3	17.4	
3	1317	robe	99.2	71.9	27.3	
4	1060	bas	100.0	39.3	60.7	
5	1050	culotte	50.7	78.4		27.7
6	1000	jupe	61.1	65.1		4.0
7	973	pantalon	78.1	49.5	28.6	
8	921	chapeau	84.6	38.9	45.7	
9	842	blouse	54.7	52.4	2.3	
10	810	gilet	12.5	80.3		67.8
11	807	chaussette	14.7	78.3		63.6
12	789	soulier	98.2	14.2	84.0	
13	742	cravate	46.2	47.6	1.4	
14	713	tablier	10.1	71.3		61.2
15	705	veste	4.0	75.0		71.0
16	656	gant	54.1	32.3	21.8	
17	602	chandail	86.9	1.9	85.0	
18	547	combinaison	8.0	54.5		46.5
19	494	pull-over	1.1	53.7		52.6
20	435	casquette	39.4	18.8	20.6	
21	410	caleçon	9.6	38.1		28.5
22	405	slip	21.7	28.6		6.9
23	370	béret	3.7	38.1		34.4
24	369	cache-nez	.1	40.7		40.6
25	344	imperméable	1.0	37.2		26.2
26	342	pyjama	30.0	15.4	14.6	
27	338	mitaines	48.5	1.3	47.2	
28	335	corsage	.0	37.0		37.0
29	307	corset	9.8	26.6		16.8
30	302	blouson	.2	33.1		32.9
31	291	pardessus	12.7	22.3		9.9
32	238	botte	29.7	4.2	25.5	
33	231	socquette	.0	25.5		25.5
34	230	ceinture	17.5	12.3	5.2	
35	226	canadienne	.0	25.0		25.0
36	213	mouchoir	24.5	5.3	19.2	
37	176	foulard	11.4	10.9	0.5	
38	174	mitaine	6.3	14.4		8.1
39	173	bonnet	10.7	11.1		0.4
40	161	chaussure	5.6	13.6		8.0
41	160	maillot	.0	17.6		17.6
42	151	tricot	.1	16.5		16.4
43	150	capuchon	1.9	15.1		13.2
44	143	jupon	18.8	1.7	17.1	
45	138	soutien-gorge	.4	14.9		14.5
46	133	brassière	17.8	1.4	16.4	
47	125	pantoufle	13.0	4.0	9.0	

48	124	bottine	18.3	.1	18.2	
49	123	cache-col	.0	13.6		13.6
50	122	shorts	6.1	8.9		2.8
51	119	costume	6.8	8.0		1.2
52	117	gabardine	.0	12.9		12.9
53	108	jaquette	13.3	1.9	11.4	
54	107	chemise de nuit	.8	11.1		10.3
55	102	flanelle	.0	11.2		11.2
56	92	habit	13.6	.0	13.6	
57	91	changes de dessous	13.5	.0	13.5	
58	82	couvre-chaussures	12.2	.0	12.2	
59	80	écharpe	.0	8.8		8.8
60	74	cape	1.1	7.3		6.2
60	74	chemisette	.0	8.1		8.1
62	71	bas de nylon	9.5	.7	8.8	
63	68	camisole	9.0	.7	8.3	
64	67	robe de chambre	7.2	1.9	5.3	
65	66	bretelle	8.3	1.1	7.2	
65	66	chausson	3.7	4.5		0.8
67	63	chemisier	.0	6.9		6.9
68	59	sweater	8.7	.0	8.7	
69	58	passe-montagne	.1	6.3		6.2
70	56	paletot	3.1	3.8	0.7	
71	55	suit	8.1	.0	8.1	
72	53	sous-vêtement	7.8	.0	7.8	
73	51	casque	6.9	.4	6.5	
74	49	habit de bain	7.2	.0	7.2	
75	46	anorak	.0	5.0		5.0
76	45	cluque	6.6	.0	6.6	
76	45	patins	6.6	.0	6.6	
76	45	scarf	6.6	.0	6.6	
76	45	tuque	6.6	.0	6.6	
80	44	lunettes	6.3	.1	6.2	
81	41	capot	6.1	.0	6.1	
81	41	sabot	.1	4.4		4.3
83	40	boucle	5.6	.2	5.4	
83	40	bouton	2.6	2.4	0.2	
83	40	col	.1	4.3		4.2
83	40	couche	5.9	.0	5.9	
83	40	crinoline	5.9	.0	5.9	
83	40	salopettes	5.9	.0	5.9	
89	39	sarreau	.0	4.3		4.3
89	39	tailleur	.0	4.3		4.3
91	37	bricole	5.5	.0	5.5	
91	37	complet	4.0	1.1	2.9	
93	36	feutres	.0	3.9		3.9
93	36	lacet	4.9	.3	4.6	
95	35	belt	5.2	.0	5.2	
95	35	slacks	5.2	.0	5.2	
97	34	soutane	4.9	.1	4.8	
98	33	bas de robe	4.9	.0	4.9	
98	33	espadrille	4.9	.0	4.9	

TABLE 1

1.2 FOOD
(The first hundred out of 588 word-types)

Rank	Population	Name (word-type)	Acadia	France	A-F = %	F-A = %
1	816	pain	83.1	67.1	16.0	
2	761	viande	77.9	61.8	16.1	
3	689	lait	91.2	19.8	71.4	
4	655	carotte	75.8	37.8	38.0	
5	538	patate	79.7	.5	79.2	
6	516	chou	61.1	27.4	33.7	
7	468	pomme	49.5	35.2	14.3	
8	458	eau	40.6	48.3		7.7
9	427	gâteau	52.6	19.0	33.6	
10	415	beurre	44.1	30.8	13.3	
11	409	café	38.9	38.3	0.6	
12	398	soupe	36.9	39.1		2.2
13	379	orange	43.8	21.9	21.9	
14	377	oeuf	47.9	14.3	33.6	
15	375	fromage	35.7	35.2	0.5	
16	372	navet	49.2	10.7	38.5	
17	340	pomme de terre	15.1	62.1		47.0
18	314	poisson	29.6	30.0		0.4
18	314	vin	1.4	79.3		77.9
20	297	tomate	41.6	4.4	37.2	
21	273	banane	32.2	14.6	17.6	
22	268	confitures	25.1	25.8		0.7
22	268	salade	16.5	40.9		24.4
24	250	thé	36.1	1.8		34.3
25	236	biscuit	34.8	.5	34.3	
26	235	pois	34.9	.0	34.9	
27	203	fève	30.0	.2	29.8	
28	194	céréale	28.8	.0	28.8	
29	193	tarte	28.5	.2	28.3	
30	181	poire	13.3	23.7		10.4
31	173	blé d'Inde	25.7	.0	25.7	
32	155	sucre	12.7	18.0		5.3
33	153	mélasse	22.7	.0	22.7	
34	150	cidre	.0	39.1		39.1
35	145	haricot	1.0	36.0		35.0
36	144	betterave	20.8	1.0	19.8	
36	144	bière	2.8	32.6		29.8
36	144	poulet	16.3	8.8	7.5	
39	142	jus de tomate	21.1	.0	21.0	
40	137	liqueur	20.3	.0	20.3	
41	131	cacao	19.4	.0	19.4	
42	128	jus d'orange	19.0	.0	19.0	
43	123	pêche	15.3	5.2	10.1	
44	116	chocolat	6.6	18.5		11.9
45	112	limonade	2.2	25.3		23.1
46	98	concombre	14.4	.2	14.2	
46	98	fraise	13.2	2.3	10.9	
46	98	jus de pomme	14.5	.0	14.5	

49	96	jus	14.2	.0	14.2	
49	96	légume	4.7	16.7		12.0
51	90	crême glacée	13.3	.0	13.0	
52	88	oignon	11.7	2.3	9.4	
53	86	jambon	9.8	5.2	4.6	
54	80	saucisse	9.0	4.9	4.1	
55	77	bonbon	10.8	1.0	9.8	
56	76	poule	11.1	.2	10.9	
57	75	sel	5.6	9.6		4.0
58	70	bacon	10.4	.0	10.4	
58	70	poireau	.2	17.7		17.5
60	69	fruits	8.7	2.6	6.1	
61	66	pamplemousse	9.8	.0	9.8	
62	65	spaghetti	9.6	.0	9.6	
63	64	macaroni	9.5	.0	9.5	
64	63	rôti	6.6	4.6	2.0	
65	62	jello	9.2	.0	9.2	
65	62	pâté	1.6	13.3		11.7
67	61	morue	8.1	1.5	6.6	
68	60	laitue	8.6	.5	8.1	
69	59	chou-fleur	1.7	12.2		10.5
69	59	naveau	8.7	.0	8.7	
69	59	riz	4.7	7.0		2.3
72	58	framboise	8.6	.0	8.6	
73	57	cerise	6.6	3.1	3.5	
73	57	homard	7.7	1.3	6.4	
75	55	céleri	7.8	.5	7.3	
75	55	petit pois	.0	14.3		14.3
77	54	poivre	3.8	7.3		3.5
78	53	bologna	7.8	.0	7.8	
79	52	sirop	7.4	.5	6.9	
80	50	prune	5.6	3.1	2.5	
81	49	lentille	.2	12.2		12.0
82	48	citron	6.6	.7	5.9	
83	47	beurre d'arachide	6.9	.0	6.9	
84	45	dessert	2.8	6.7		3.9
84	45	gruau	6.6	.0	6.6	
86	44	bleuet	6.5	.0	6.5	
86	44	crême	5.6	1.5	4.1	
88	43	hareng	5.9	.7	5.2	
88	43	radis	4.7	2.8	1.9	
90	39	nouille	.0	10.1		10.1
91	38	citrouille	5.5	.2	5.3	
92	36	dinde	5.3	.0	5.3	
92	36	noix	2.8	4.4		1.6
94	34	corn flakes	5.0	.0	5.0	
94	34	crêpe	5.0	.0	5.0	
94	34	sardine	2.0	5.2		3.2
94	34	steak	5.0	.0	5.0	
98	33	maquereaux	4.4	.7	3.7	
99	32	bean	4.7	.0	4.7	
99	32	sauce	4.6	.2	4.4	

TABLE 1

1.3 TRANSPORTATION
(The first hundred out of 958 word-types)

Rank	Population	Name (word-type)	Acadia	France	A-F = %	F-A = %
1	1378	train	90.8	92.1		1.3
2	1352	avion	85.7	92.8		7.1
3	1270	camion	74.8	91.2		16.4
4	1260	bateau	89.7	79.5	10.2	
5	862	auto	59.6	55.6	4.0	
6	840	autobus	70.7	45.2	25.5	
7	781	bicyclette	59.0	46.8	12.2	
8	751	cheval	78.2	29.8	48.4	
9	655	moto	.3	74.2		73.9
10	527	charrette	8.3	54.0		45.7
11	502	tracteur	38.0	30.0	8.0	
12	454	hélicoptère	40.0	23.2	16.8	
13	449	car	2.5	49.2		46.7
14	431	vélomoteur	.0	49.0		49.0
15	410	automobile	34.5	22.0	12.5	
16	398	voiture	15.2	34.4		19.2
17	368	vélo	.0	41.8		41.8
18	362	traîneau	48.0	7.0	41.0	
19	350	micheline	.0	39.8		39.8
20	327	brouette	11.5	29.0		17.5
21	317	camionnette	2.2	34.4		32.2
22	315	taxi	18.8	22.4		3.6
23	226	autocar	.6	25.2		24.6
24	225	paquebot	.9	24.9		24.0
25	215	remorque	.1	24.3		24.2
26	207	transport	32.9	.1	32.8	
27	205	métro	.4	22.9		22.5
28	195	péniche	.0	22.1		22.1
29	189	scooter	1.9	20.1		18.2
30	185	barque	.3	20.8		20.5
31	184	canot	23.5	4.2	19.3	
31	184	tramway	4.0	18.0		14.0
33	174	diligence	.1	19.6		19.5
34	172	ski	25.9	1.1	24.8	
34	172	wagon	4.6	16.2		11.6
36	164	autorail	.0	18.6		18.6
37	156	chariot	.6	17.2		16.6
38	155	âne	17.1	5.4	11.7	
39	149	chien	23.0	.5	22.5	
40	147	à pied	22.2	.9	21.3	
41	145	sous-marin	13.4	6.9	6.5	
42	136	chameau	14.5	5.1	9.4	
43	129	jet	20.6	.0	20.6	
44	126	express	.4	13.9		13.5
45	124	truck	19.8	.0	19.8	
46	117	raquette	18.7	.0	18.7	
47	111	carrosse	9.4	5.9	3.5	
48	106	fusée	15.2	1.2	14.0	

49	104	char	7.5	6.4	1.1	
50	94	jeep	9.7	3.7	6.0	
51	93	locomotive	.4	10.2		9.8
52	92	motocyclette	7.8	4.8	3.0	
53	91	mobylette	.0	10.3		10.3
54	89	charrue	12.6	1.1	11.5	
55	87	boeuf	10.0	2.7	7.3	
55	87	téléphérique	.0	9.8		9.8
57	83	omnibus	.0	9.4		9.4
58	79	poussette	.0	8.9		8.9
59	71	tram	.0	8.0		8.0
60	65	éléphant	8.4	1.3	7.1	
60	65	hydravion	.0	7.3		7.3
62	63	vedette	.0	7.1		7.1
63	62	aéroplane	5.7	2.9	2.8	
64	56	cargo	.4	6.0		5.6
65	55	ski-mobile	8.8	.0	8.8	
65	55	traîne	8.8	.0	8.8	
67	53	ambulance	4.3	2.9	1.4	
67	53	side-car	.0	6.0		6.0
67	53	tombereau	.0	6.0		6.0
68	49	bicycle	7.8	.0	7.8	
68	49	carriole	4.3	2.5	1.8	
70	47	roulotte	1.1	4.5		3.4
70	47	snow-mobile	7.5	.0	7.5	
72	46	tank	1.1	4.4		3.3
73	43	vanne	6.8	.0	6.8	
74	42	bicycle à gaz	6.7	.0	6.7	
75	41	toboggan	6.5	.0	6.5	
76	38	navire	2.8	2.2	0.6	
77	37	bicyclette à gaz	5.9	.0	5.9	
78	33	avion à réaction	1.2	2.8		1.6
79	32	poney	4.9	.1	4.8	
79	32	tandem	.0	3.6		3.6
80	31	tricycle	3.3	1.1	2.2	
81	30	bus	4.6	.1	4.5	
81	30	chemin de fer	.8	2.8		2.0
81	30	marche	4.1	.4	3.7	
81	30	rapide	.0	3.4		3.4
85	29	bateau à voile	4.6	.0	4.6	
86	28	charreton	.0	3.1		3.1
86	28	trailer	4.4	.0	4.4	
88	27	ballon	1.4	2.0		0.6
89	26	aveneur	.4	2.6		2.2
89	26	motor cycle	4.1	.0	4.1	
91	25	citram	.0	2.8		2.8
91	25	soucoupe volante	3.5	.3	3.2	
93	24	ferry	3.8	.0	3.8	
93	24	funiculaire	.0	2.7		2.7
95	23	trolley	3.2	.3	2.9	
96	22	poste	.1	2.3		2.2
96	22	pousse-pousse	.9	1.8		0.9
98	21	bateau à vapeur	1.2	1.4		0.2

TABLE 1

1.4 PASTIMES
(The first hundred out of 1328 word-types)

Rank	Population	Name (word-type)	Acadia	France	A-F = %	F-A = %
1	477	carte	48.0	48.8		0.8
2	402	football	27.1	66.3		39.2
3	380	balle	46.6	23.0	23.0	
4	366	course	25.1	59.6		34.5
5	353	dame	32.8	41.2		8.4
6	279	hockey	42.4	1.4	41.0	
7	264	cinéma	2.3	72.8		70.5
8	236	ballon	23.1	25.4		2.3
9	224	tennis	21.0	25.7		4.7
10	221	théâtre	.9	62.8		61.9
11	213	bille	16.8	30.4		13.6
12	189	cache-cache	17.2	22.8		5.6
13	187	base-ball	28.9	.0	28.9	
14	182	cachette	18.9	17.5	1.4	
15	176	quilles	24.1	5.8	18.3	
16	169	poupée	13.6	23.6		10.0
17	157	corde	14.7	18.1		3.4
18	154	bingo	23.8	.0	23.8	
19	153	ski	21.3	4.3	17.0	
20	151	ping-pong	11.7	21.9		10.2
21	145	chat	22.0	.8	21.2	
21	145	domino	8.5	26.3		17.8
23	143	télévision	20.7	2.6	18.1	
24	136	ballon-panier	21.0	.0	21.0	
25	132	oie	.0	38.5		38.5
26	128	basket-ball	11.3	16.0		4.7
27	126	monopoly	16.7	5.2	11.5	
28	113	boxe	2.9	27.4		24.5
29	109	lecture	9.6	13.7		4.1
30	108	nage	9.9	12.8		2.9
31	107	promenade	5.5	20.7		15.2
32	104	bal	.1	30.1		30.0
33	102	bicyclette	12.4	6.4	6.0	
34	101	nain-jaune	.0	29.5		29.5
35	100	pêche	10.0	10.2		0.2
36	93	cow-boy	14.2	.0	14.2	
36	93	radio	13.4	1.4	12.0	
38	91	patinage	13.4	1.1	12.3	
39	88	drapeau	13.6	.0	13.6	
39	88	loto	.0	25.7		25.7
41	87	rugby	.0	25.4		25.4
42	86	cirque	1.8	21.6		19.8
43	82	livre	11.1	2.9	8.2	
43	82	mouchoir	10.3	4.3	6.0	
45	78	école	10.5	2.9	7.6	
46	76	tag	11.7	.0	11.7	
47	70	golf	10.5	.5	10.0	
48	67	chasse	6.5	7.3		0.8

48	67	musique	9.1	2.3	6.8	
50	66	dada	.0	19.2		19.2
50	66	tape	1.7	16.0		14.3
52	65	ballon-prisonnier	10.0	.0	10.0	
52	65	danse	8.6	2.6	6.0	
54	63	catch	.1	18.1		18.0
54	63	ronde	.3	17.8		17.5
56	62	badminton	9.6	.0	9.6	
57	61	lire	8.8	1.1	7.7	
58	59	serpent	9.1	.0	9.1	
59	57	piano	8.5	.5	8.0	
60	56	marche	4.4	7.8		3.4
61	53	dessin	7.1	2.0	5.1	
62	52	baigner	8.0	.0	8.0	
62	52	gendarme-voleur	.0	15.2		15.2
62	52	petits chevaux	.0	15.2		15.2
65	51	patiner	7.9	.0	7.9	
65	51	volleyball	2.3	10.5		8.2
67	50	femme	7.7	.0	7.7	
68	49	ours	6.5	2.0	4.5	
69	48	balançoire	4.6	5.2		0.6
69	48	neige	7.4	.0	7.4	
71	47	dessiner	6.8	.8	6.0	
71	47	tic-tac-to	7.2	.0	7.2	
73	46	auto	5.7	2.6	3.1	
73	46	boule	5.4	3.2	2.2	
73	46	chat-perché	.0	13.4		13.4
73	46	madame	7.1	.0	7.1	
73	46	patin	6.2	1.7	4.5	
78	45	puzzle	6.9	.0	6.9	
79	44	lutte	6.3	.8	5.5	
80	43	baskett	.0	12.5		12.5
80	43	hand-ball	.0	12.5		12.5
80	43	traîneau	6.6	.0	6.6	
83	42	train	5.5	1.7	3.8	
84	41	gouret	6.3	.0	6.3	
85	40	bateau	4.9	2.3	2.6	
85	40	bois	6.0	.2	5.8	
85	40	loup	1.8	8.1		6.3
88	39	chant	5.8	.2	5.6	
88	39	danse	6.0	.0	6.0	
88	39	écrire	5.7	.5	5.2	
88	39	piano	.1	11.1		11.0
92	38	cheval	4.1	3.2	0.9	
92	38	corde à danser	5.8	.0	5.8	
94	37	échecs	4.4	2.3	2.1	
94	37	histoire	5.2	.8	4.4	
94	37	maison	5.7	.0	5.7	
97	36	casse-tête	5.5	.0	5.5	
97	36	chanter	5.5	.0	5.5	
99	35	camion	5.2	.2	5.0	
99	35	colin-maillard	.9	8.4		7.5
99	35	soft-ball	5.4	.0	5.4	

TABLE 2

CATEGORIZATION

		Acadia	France	
1. CONCEPTUAL FIELD: Clothing		49.8%	50.2%	
2. CONCEPT CATEGORY: Head-dress		53.4%	46.6%	
3. CATEGORY SEGMENTS:				
	hat	84.7%	42.8%	
	cap	39.5%	19.7%	
	beret	3.7%	38.1%	
	helmet	7.1%	.4%	
	bonnet	17.3%	11.1%	
	hood	3.6%	17.7%	
	kerchief	.1%	.0%	
4. NAMES:				Rank
	chapeau	84.6%	38.9%	8
	feutre	.0%	3.9%	93
	hat	.1%	.0%	307
	casquette	39.4%	18.8%	20
	calotte	.1%	.2%	223
	képi	.0%	.7%	181
	béret	3.7%	38.1%	23
	casque	6.9%	.4%	73
	helmet	.2%	.0%	257
	bonnet	10.7%	11.1%	39
	tuque	6.6%	.0%	76
	passe-montagne	.1%	6.3%	69
	capuchon	1.9%	15.1%	43
	capuche	1.4%	2.3%	102
	capeline	.0%	.1%	307
	chaperon	.1%	.0%	307
	béguin	.0%	.2%	257
	hood	.2%	.0%	257
	coiffe	.1%	.0%	307

TABLE 3

DIVERSITY

	Native words		Loans		Proper names		Neologisms	
	Acadia	France	Acadia	France	Acadia	France	Acadia	France
3.1 Clothing								
Occasional wear	5	2						
Foot-wear	17	16	12					
Suits	4	3	2					
Night wear	7	4	1					
Slacks and trousers	2	3	6					
Swim-wear	5	4	3	1				
Accessories	37	23	11					
Dresses	4	6	4					
Head dress	19	14	5					
Stocking	5	6	2					
Shirts	6	11	6	1				
Underwear	19	13	7	1				
Top-coats	13	19	8					
Jackets	2	5	2	1				
Aprons	3	7	2					
3.2 Food								
Tinned food	1	0	3		2			
Juices	15	1	2		1			
Bakery products	4	1	2					
Cereals	6	1	5		5			
Syrups	4	1	2					
Desserts	6	4	4		1			
Beverages	8	4	2		3			
Fats	3	1						
Poultry	9	8	1					
Meat products	9	7	5					
Vegetables	36	28	9					1
Milk products	9	7	5		4			
Fruits	26	19	10					
Fish	12	14	1					
Breads	4	1	2					
Sweets	11	3	4					
Meats	22	16	9	1				
Jams	4	1	1					
Crustaceans	3	7						
Soups	2	2						
Food pastes	5	2						
Condiments	14	11	5		2			
Sugars	4	1						
Non alcoholic drinks	4	8			4	2		
Starches	5	2	2					
Shellfish	1	6						
Alcoholic drinks	4	8	3	2		4		

3.3 Transportation

Snow vehicles	5	2	6	1	1		
Riding (animals)	15	15	2				
Lifts	2	1	1				
Air travel	6	16	2		3	1	
Hauling	6	9	4			1	
Automobiles	10	8	3		8	9	
Hand vehicles	1	5					
Sea travel	21	49	10	1			
Bicycles	7	7					
Public transportation	4	5	3	5		1	
Rail travel	4	14	1	1	2	1	
Motorcycles	5	6	2	1		3	1
Farm vehicles	3	5					
Military vehicles	2	5					
Horse-drawn vehicles	2	6					
Cable cars		2					

3.4 Pastimes

Golf			1	1			
Winter games	18	5	8				
Music	16	7	1				
Bowling and billiards	2	2	3				
Camping	4	3	2				
Play-action	64	25	10	1			
Reading	11	6	2				
Manual hobbies	30	10	3			7	
Water sports	11	6	2	1		6	
Ball games	9	9	4	2			
Physical skills	20	11	4	2			
Motor sports	7	8	1	1			
Games of strategy	12	6	8			1	1
Hunting and fishing	5	2					
Yard games	37	44	12			3	
Racket (net) sports	2	2	5	5			
Spectator entertainment	14	8	5				
Field games	1	1	6	7			
Dancing	4	3	5			1	
Games of chance	10	19	10	3		2	
Track sports	6	10	5			3	
Gymnastics	4	8	2				
Dolls games	6	3	2				
Cycling, riding & hiking	9	8				1	
Boxing and wrestling	5	5	3			1	
Holiday games	3	12	3			1	

TABLE 4

SELECTED ITEMS

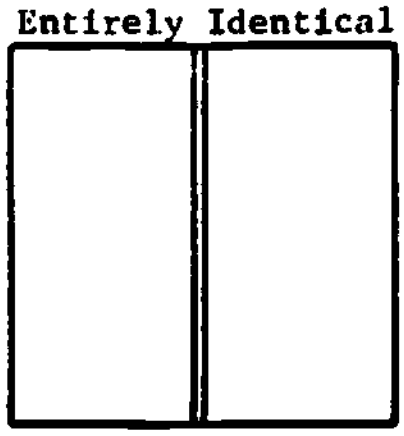
Concept	Acadia %	France %
wine	1.4	79.3
beret	3.7	38.1
smock	.8	52.7
metro	.7	23.4
soccer	8.5	26.3
hockey	42.0	1.4
sleigh	54.5	7.3
skate	29.2	2.0
juice	74.6	.2
bingo	87.0	0.0

Selected items from four conceptual fields.

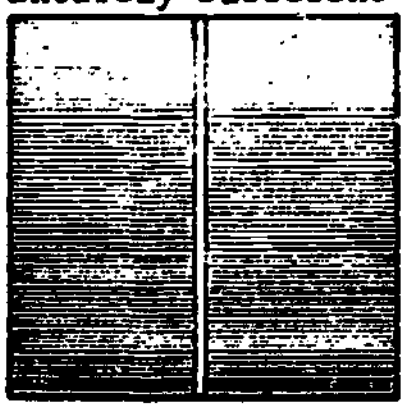
Figure 1

STATISTICAL PATTERNS OF CULTURE DIFFERENCE

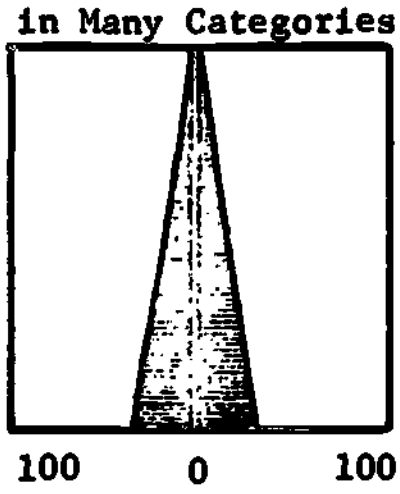
Entirely Identical



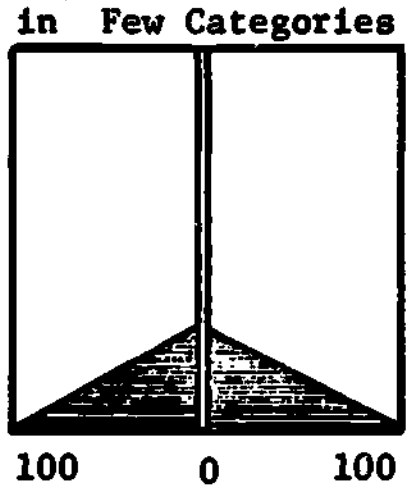
Entirely Different



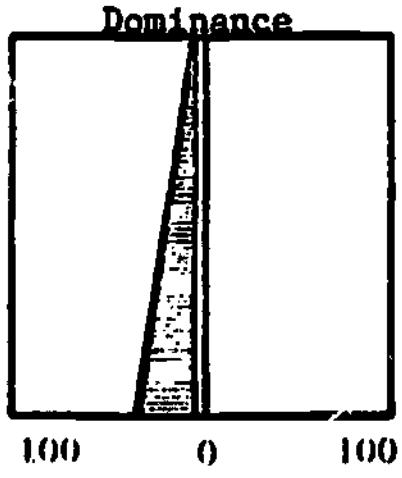
Minor Differences in Many Categories



Major Differences in Few Categories



Diffuse Minor Dominance



Concentrated Major Dominance

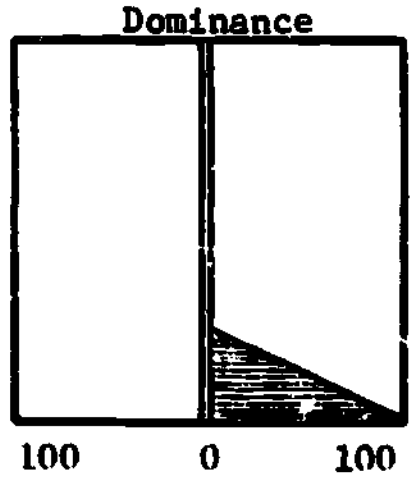
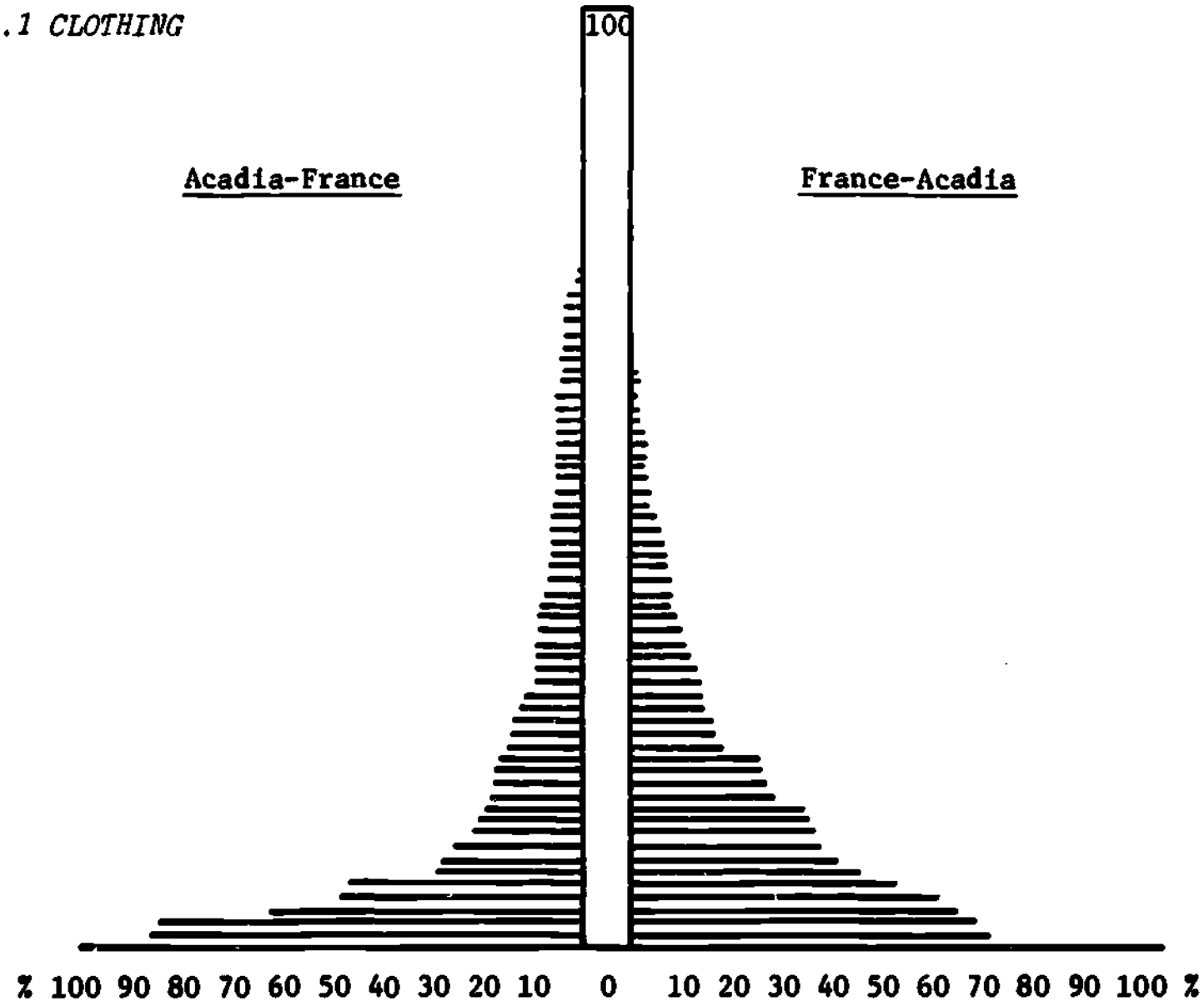
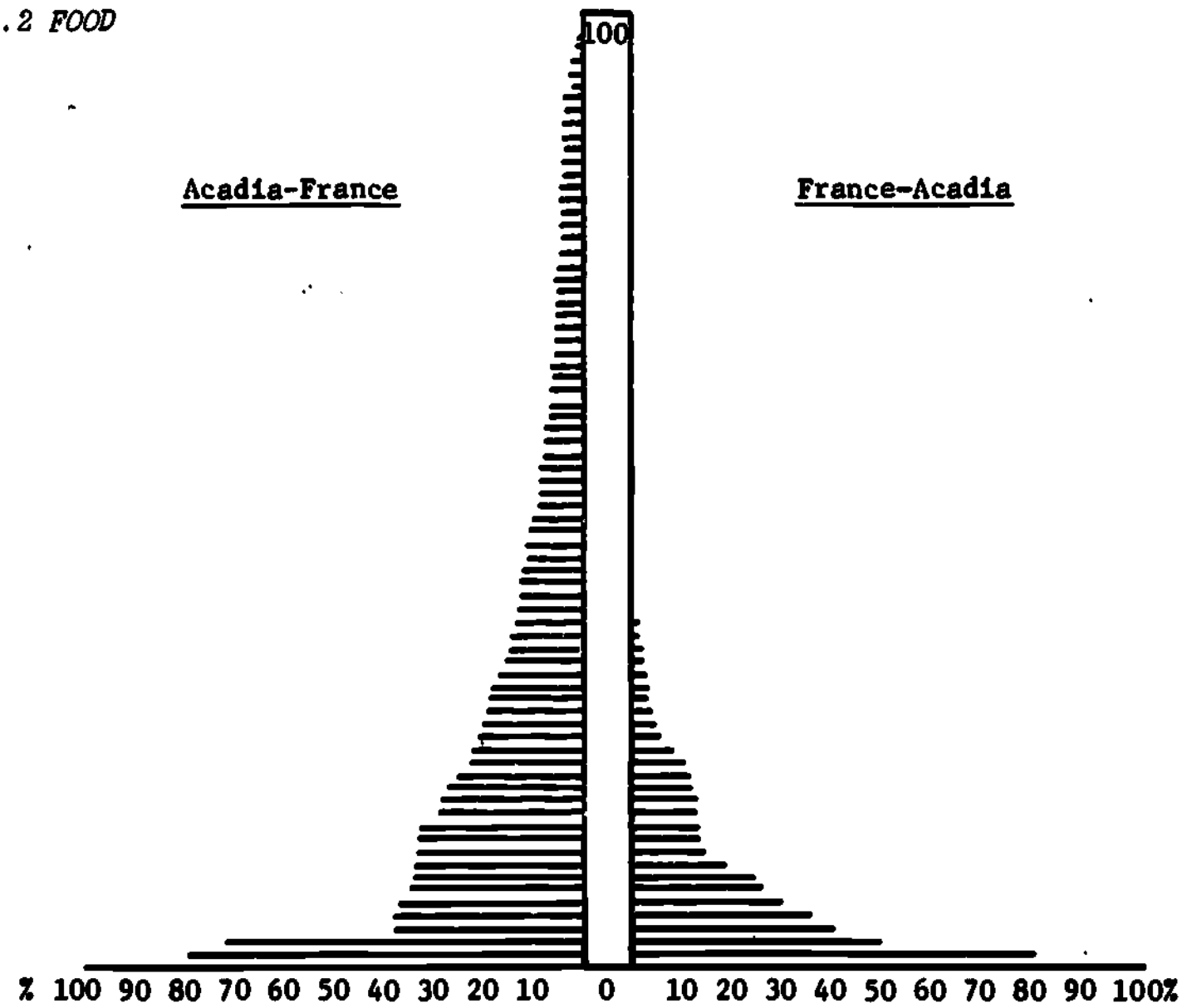


Figure 2

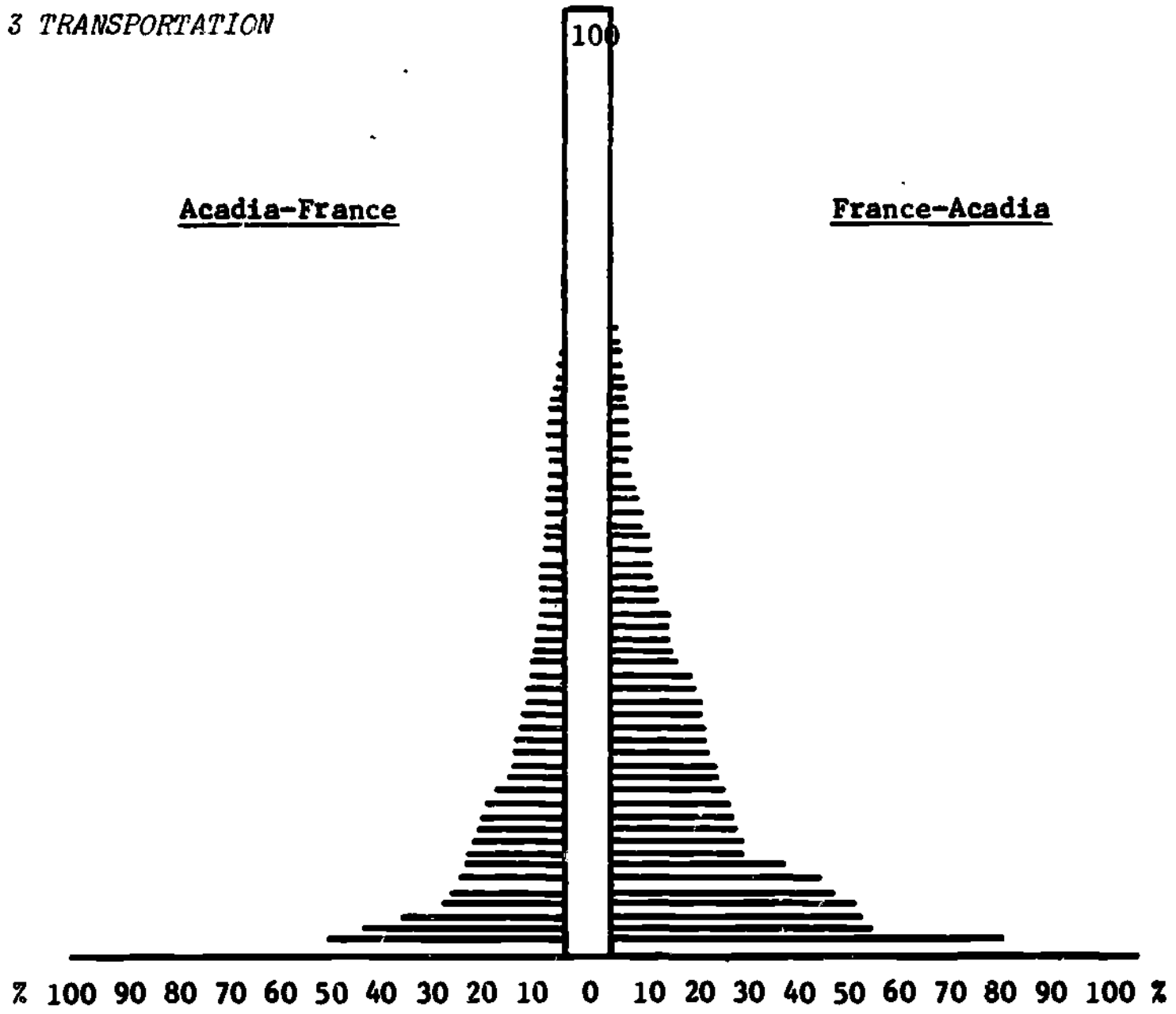
2.1 CLOTHING



2.2 FOOD



2.3 TRANSPORTATION



2.4 PASTIMES

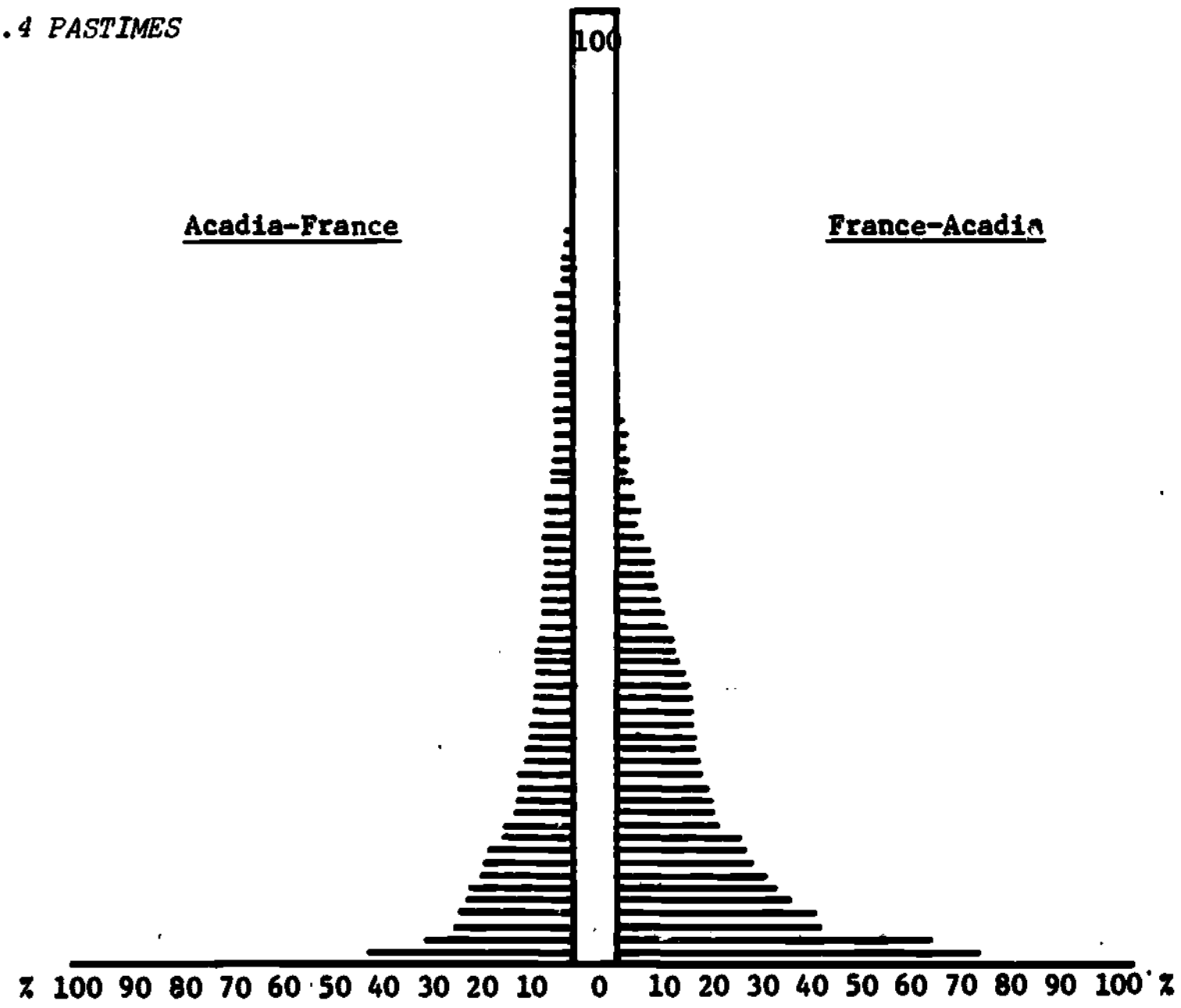
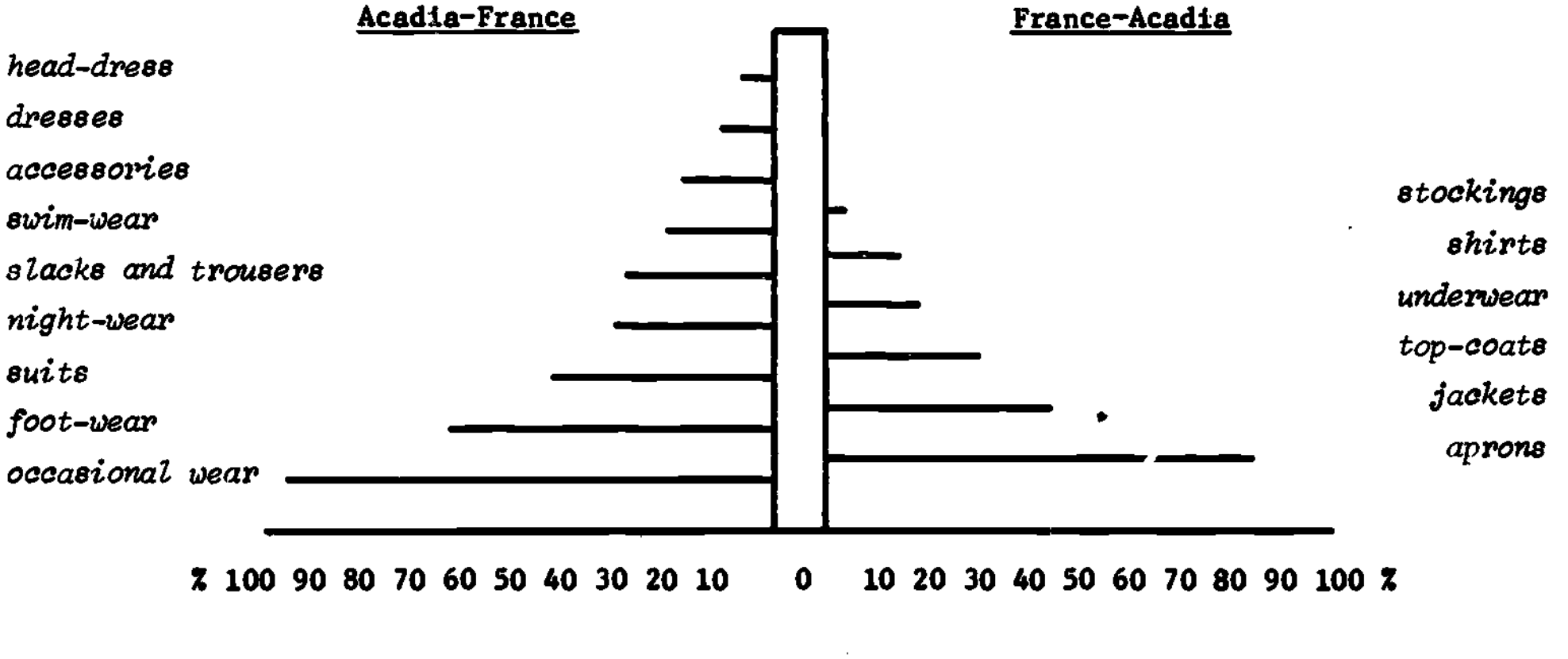
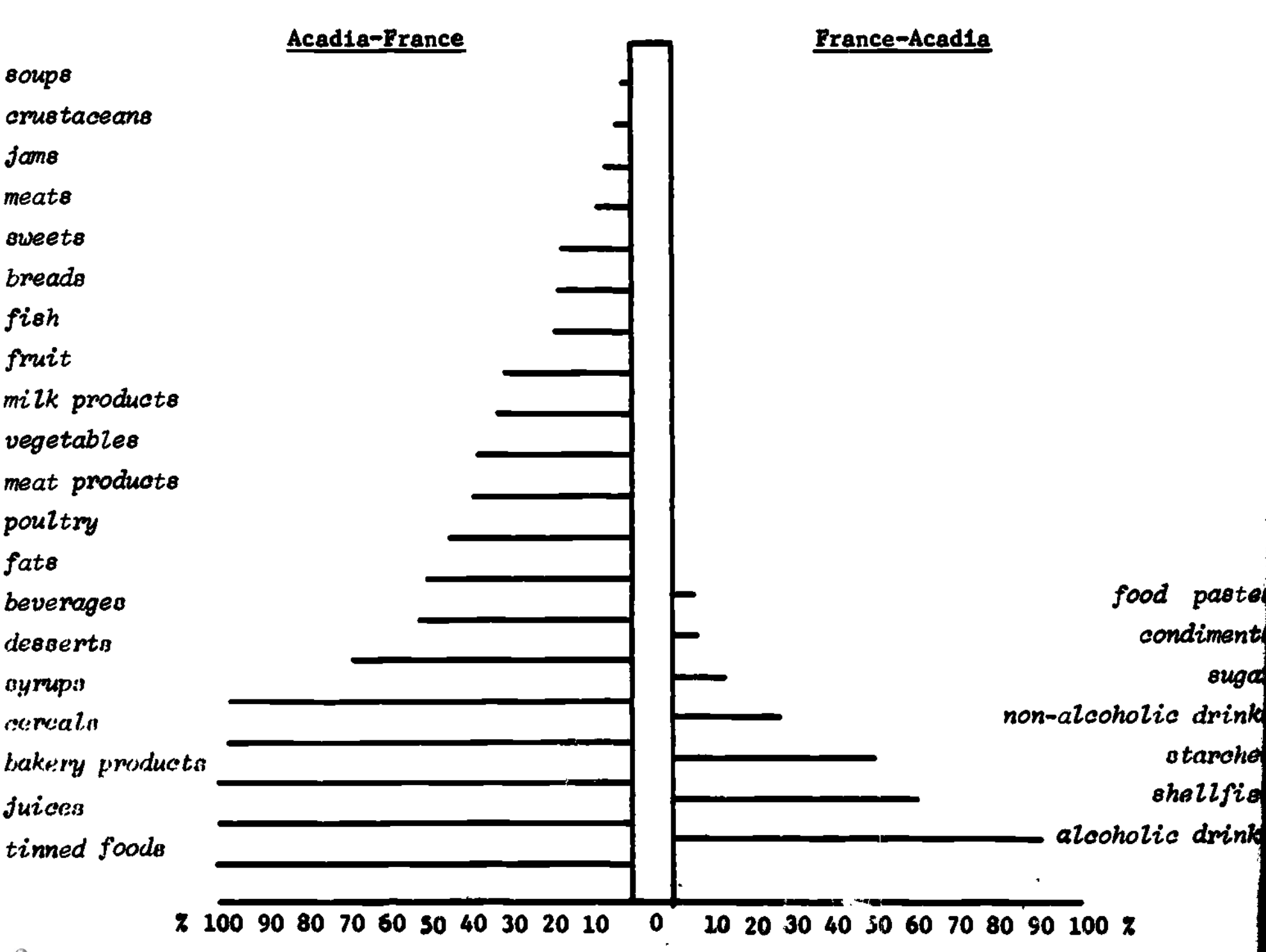


Figure 3

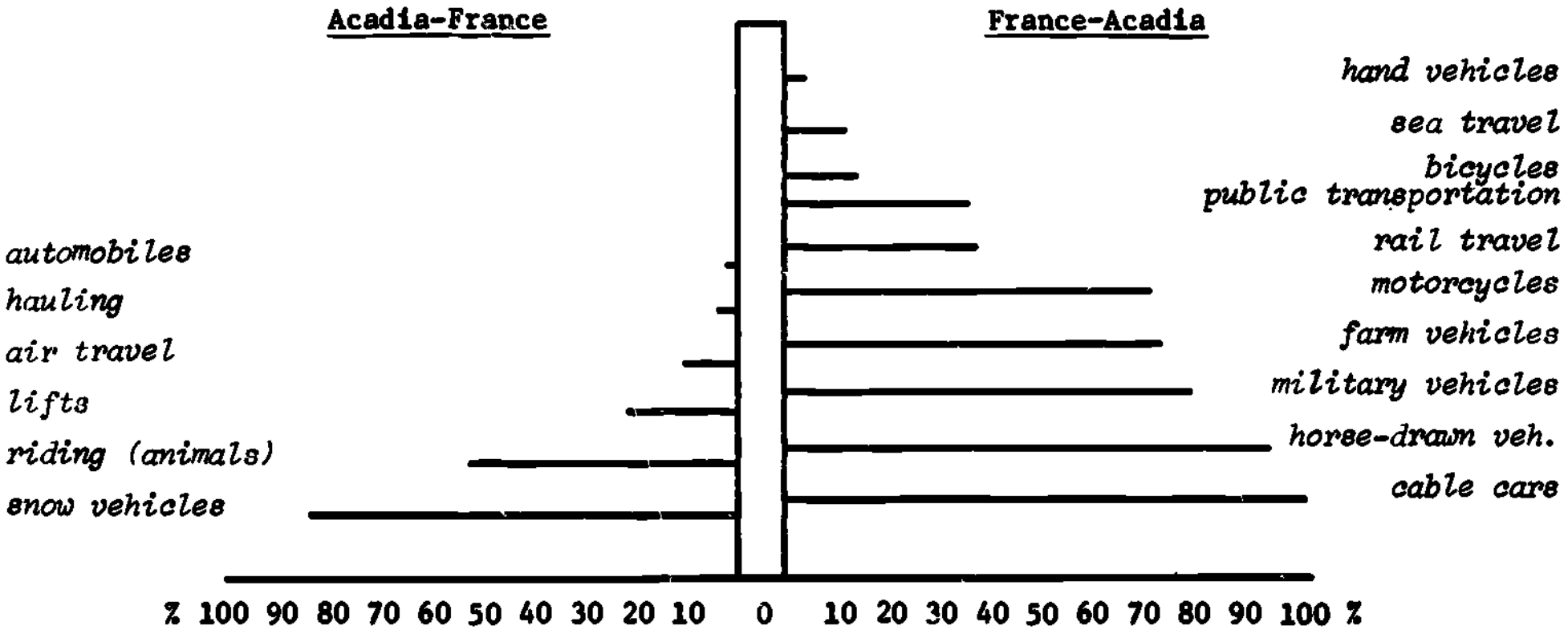
3.1 CLOTHING



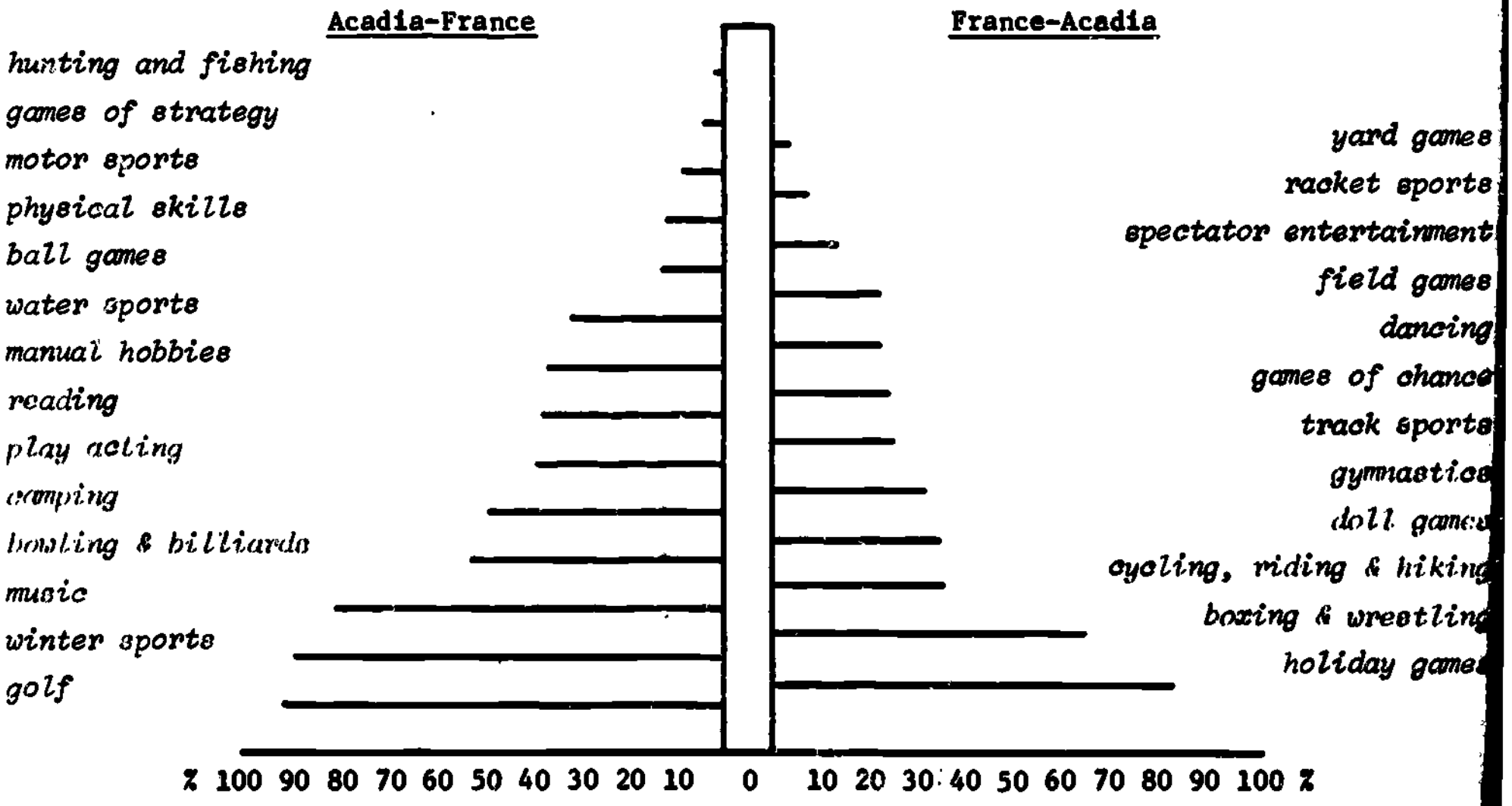
3.2 FOOD



3.3 TRANSPORTATION



4.3 PASTIMES



Acknowledgments: Of the persons who helped obtain and compile the data, special thanks are due to Michel Dabène, Jean-Guy Savard, Yane Godbout, Faulette Weller, Monique Daigle, the Crédif and the Acadian Teachers Association. Thanks are also due to the National Museum, the Canada Council, the Office de la langue française and the Centre de traitement de l'information of Laval University for material support for field work, analysis and computation.