Slices of Meaning:
A Study on Nouns of Portions

Salvador Climent
Maria Antònia Martí

Report LSI-95-32-R
SLICES OF MEANING: A STUDY ON NOUNS OF PORTIONS.

Salvador Climent
Universitat Politècnica de Catalunya
climent@goliat.upc.es

Maria Antònia Martí
Universitat de Barcelona
amarti@lsi.upc.es

June 1995*

ABSTRACT

Within the framework of the analysis of Part-Whole relations the case of nouns of portions (such as "slice") is studied. In preliminary sections, a corpus of Spanish of these kind of words is displayed, and several approaches to the issue are summarised. Then, syntactic and semantic analyses are set which lead to posit that Portion Nouns are not common nouns but complex grammatical objects sharing properties of nouns, specifiers, operators and predicats thus they might no longer be treated as nouns but as a special class of operators on NPs. Consequently, Portion-Whole relation should not be treated as other types of Part-Whole relations -usually represented in Lexicons by means of links between nouns. Meaning components of Portion Nouns are discussed, resulting on a general taxonomy of portion terms grounded in cognitive distinctions. Finally, a representation of the most representative cases, making use of concepts coming from Cognitive Grammar, Ray Jackendoff's Conceptual Semantics, and Pustejovsky's Theory of Qualia is outlined.

* This research has been made within the framework of the Acquilex-II Project (ESPRIT BRA 7315) which is aimed to the building of a large Lexical Knowledge Base for Natural Language Processing.
# CONTENTS

1 **Introduction** .......................................................... 1

2 **Some approaches to the Part-Whole relation** ................. 3
   2.1 Theoretical accounts ........................................... 4
      2.1.1 Winston *et al.* 1987 .................................. 4
      2.1.2 Chaffin *et al.* 1988 .................................. 5
      2.1.3 Iris *et al.* 1988 ........................................ 6
      2.1.4 Jackendoff 1991 ........................................... 8

2.2 Computational lexicons ............................................ 9
   2.2.1 EDR and WordNet ........................................... 9
   2.2.2 Acquilex LKB (Vossen & Copestake 1994) ................. 9

3 **Syntax: Are Portion Nouns nouns as the rest?** ............... 11
   3.1 Relational approaches don't suit for portions .............. 14

4 **Some remarks on [Vossen & Copestake 1994]'s approach** .... 16

5 **Semantics: On the complex behaviour of Portion Nouns** .... 18

6 **Generalisations: outlining a taxonomy of portion terms** ... 20

7 **Features for these general types** ................................ 22
   7.1 About types .................................................. 22
   7.2 About features ................................................. 23

8 **Outlining lexical entries for portion terms** .................. 26
   8.1 Examples ...................................................... 28

9 **Relevance for Natural Language Processing** .................. 33

10 **A special case: portions of abstract concepts** ................ 34

11 **Conclusions** ..................................................... 35

12 **Further work** ..................................................... 36

**References** ............................................................ 38
INTRODUCTION

1. Our purpose here is characterising the semantics and approaching a lexical representation of words such as English slice, portion or sheet or Spanish rebanada lámina or trozo\(^0\). Intuitively, we are dealing with nouns denoting non-functional parts of objects or masses; probably obtained by some sort of detachment process; and which although they can be considered as a new object distinct from the whole, they bear traits of their own such as form or others. Moreover, they usually need further syntagmatic (or contextual) specification\(^1\), namely a modifier in genitive, to indicate of which kind of object or material they are portions of.

Here is a subset of the Spanish lexicon, most of them extracted from the dictionary by searching for definitions containing words such as pedazo (piece, bit), porción (portion) or parte (part), to get a first outlook on this kind of words:

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{Spanish} & \textbf{English} \\
\hline
rebanada & slice \\
lámina & sheet \\
trozo & piece, bit, portion \\
porción & portion \\
parte & part \\
\hline
\end{tabular}
\end{table}

\(^0\) Although the discussion is written in English and intended to be language-independently general for portion words, our work will focus on Spanish.

\(^1\) Usually but not always. Certainly one would not say

\begin{itemize}
\item \textit{Tengo un trozo en mi casa}
\item [I have a portion at home]
\end{itemize}

unless the hearer could pragmatically infer a portion of what we are talking about. Nevertheless, if one says

\begin{itemize}
\item \textit{Me comería un filete}
\item [I feel like eating a special-portion-of-meat]
\end{itemize}

in absence of further information the hearer will infer we are talking about a portion of beef -although a filete might also be the same part of another animal. Undoubtedly some kind of notion of prototypicity or default value is playing a role here.

\(^2\) In square brackets there is the translation to English provided by [VOX-HARRAP'S 1993] and in brackets our
andrajo [rag, tatter] (a piece torn of an old cloth or dress)
ascua [red hot coal] (a bit of any material which is burning without a flame)
añojal [-] (a stretch of land which is left unfarmed for a period of time)
bancal[-] (a rectangular stretch of land ready for sowing or farming it)
barra [bar] (rigid, prismatic or cylindrical piece which is much more longer than thick)
bocado [mouthful] (a portion of food which can be naturally held in the mouth)
brizna [bit, piece, blade] (thread-like or very thin part of anything)
cacho [piece, bit] (small bit of anything)
cala [slice] (a portion of a fruit which is cut to be tasted)
capa [coat] (uniform extension of a substance which covers something)
cara [side] (side, surface of anything)
cardambano [icicle] (long pointed piece of ice)
celda [cell] (compartment of a statistic diagram formed by a column and a horizontal line which crosses it)
cilindro [cylinder] (solid limited by a cylindrical surface and two parallel planes which form its bases)
corteza [bark, crust, peel, skin, rind] (hard and exterior part of some things such as lemons, cheese, bread, etc.)
cuadrícula [-] (set of the squares which result from perpendicular intersection of two series of parallel lines)
cuarto [quarter] (part which together with other three equal parts makes a whole)
cubo [cube] (regular solid limited by six equal squares)
cucharada [spoonful] (portion which holds in a spoon)
cuscurro [crust of bread] (end of a bread)
escalope [escalope] (thin slice of beef to be eaten either fried or soaked in bread mass and thus baked)
esquina [corner] (edge, esp. that resulting from the meeting of the walls of a building)
extremo [extreme, end] (each of the two points, degrees, moments, etc. of a thing which are the most possibly distant one from the other)
filete [sirloin] (small thin slice of either lean meat or bone-clean fish)
fragmento [fragment, piece] (piece split of something)
gajo [section] (inner division of some fruits, e.g. orange)
gota [drop] (globule of any liquid; small quantity of something)
grano [grain] (small rounded bit of any substance)
hoja [sheet, leaf] (thin sheet of any material)
jarra [jug, mug] (container with a handle and a wide opening at the top)
jira [strip of cloth] (quite large and long bit cut or torn of a cloth)
jirón [shred, strip] (a bit torn of a cloth)
lágrima [tear] (drop of the humour segregated by the lachrymal gland which is dropped by the eyes; drop of the humour segregated by certain trees after being pruned; very small portion of any liquor)
lámina [sheet] (thin sheet of material, esp. metal)
lascal[-] (a small thin chunk detached from a stone)
lingua [strip] (thing which more or less has a form like a tongue)
 línea [line] (end, limit; de agua: intersection between the free surface of a liquid and that of the body which floats in it)
lingote [ingot] (piece or bar of rough metal)
listón [lath] (narrow piece of board)
loncha [slice] (lonja)
**lonja** [slice, rasher] (long wide and not very thick portion which has been cut off or detached from something, e.g. ham)

**mechón** [lock] (portion of hair or thread)

**mendrugo** [hard crust] (portion of stale bread to be cast aside)

**miga, migaja** [crumb] (small portion of anything, e.g. bread)

**mitad** [half] (each one of the two equal parts in which a whole is divided)

**morro** [knob, round end, snout] (snout of animals; round portion similar in shape to a head; end of a jetty)

**oblea** [-] (very thin, usually circular, leaf made of flour and water baked in a mould; plaque cut off of an ingot)

**orejón** [-] (bit of a peach or another fruit, dried at the air and the sun)

**parcela** [plot, portion] (small portion of land)

**pedazo** [piece, bit] (part of a thing detached from the whole)

**película** [film] (thin fragile membranous skin or cover)

**pie** [bottom, base, stand] (base or part in which anything leans on)

**pizca** [bit, jot, whit] (very small portion of anything)

**placa** [plaque] (sheet, plate or film formed on or put on the top of an object)

**plancha** [plate, sheet] (flat metal sheet which is thin with respect to its size)

**punta** [tip, point] (acute end of anything)

**puñado** [handful] (portion of anything which can be held in the hand)

**raja** [slice] (portion cut along or across a melon, a watermelon, cheese, etc.)

**rebanada** [slice] (long thin wide slice of something, esp. bread)

**región** [region] (portion of a territory determined by special circumstances; determined space of the body)

**rodaja** [slice] (flat round piece cut off of something, e.g. a tree or an orange)

**redondo** [lobspice] (portion of beef -of the back thigh of the animal)

**sección** [section] (cut, division; figure resulting from the intersection of a surface or a solid with another surface; figure resulting of cutting a body by a plane; part or distinct group in which a continuous whole or a set of things is divided)

**segmento** [segment] (part or portion cut of anything, line)

**solomillo** [sirloin] (in slaughterhouse animals, the muscular coat which extends between the ribs and the back)

**taco** [cube, plug, stopper] (short thick piece of wood, metal or another material)

**tercio** [third] (part which with other two equal parts it makes a whole)

**terrón** [lump] (small tight mass of soil or other substances, e.g. sugar)

**tira** [strip] (long narrow bit of a thin thing)

**trozo** [piece, chunk] (a portion of something considered apart from the rest)

**vaso** [glass] (liquid which holds in a glass)

**zurullo** [-] (chubby portion of soft material)

---

**SOME APPROACHES TO THE PART-WHOLE RELATION.**

2. Lexical networking is a point of major importance for natural language processing as long as it allows for carriage of semantic information between words and thus for meaning inference during discourse processing. Relational models of the lexicon make use of different semantic relations such as synonymy, antonymy, class-inclusion (IS_A) or meronymy (PART_OF). Among these, IS_A is usually the privileged relation, as it is regarded as permitting inheritance of properties along words, namely taxonomic organisation -thus avoiding redundancy of
semantic information when building a lexicon.

This is far not so clear about meronymy. In fact ([WINSTON et al. 1987],[IRIS et al. 1988]), transitivity is one of the major points of controversy when discussing this relation -its hypothetical status of primitiveness ([WINSTON et al. 1987],[CHAFFIN & HERRMAN 1988],[JACKENDOFF 1991]) is the other one.

These theoretical accounts, as well as some Computational Lexicology approaches to meronymy, are summarised below.

2.1 THEORETICAL ACCOUNTS.

2.1.1 WINSTON et al. 1987. The authors demonstrate that (i) meronymy is not a single relation but a family of them and (ii) it is a class of relations distinct from spatial inclusion and class inclusion. They also show that transitivity does not hold across different types of meronymic relations. This way, paradoxical cases of intransitivity previously pointed out by [LYONS 1977] and [CRUSE 1979] (cf. [IRIS et al. 1988]) as in (1) become explained by the fact that (1a) and (1b) hold different types of a PART-OF relation, namely Component-Integral Object (1a) and Member-Collection (1b):

(1)  
 a. Simpson's finger is part of Simpson
 b. Simpson is part of the Philosophy Department
 c. *Simpson's finger is part of the Philosophy Department

Six types of meronymic relations are posed:

(TABLE 1)

| (a) Component-Integral Object (pedal-bike) | +F -H +S |
| (b) Member-Collection (ship-fleet)          | -F -H +S |
| (c) Portion-Mass (slice-pie)               | -F +H +S |
| (d) Stuff-Object (steel-car)               | -F -H -S |
| (e) Feature-Activity (paying-shopping)     | +F -H -S |
| (f) Place-Area (oasis-desert)              | -F +H -S |

[WINSTON et al. 1987] also suggest that each type itself is not a unitary entity or primitive, but they are decomposable into more basic elements. They account for three of them: Functional ('"parts are/are not in a specific spatial/temporal position with respect to each other which supports their functional role with respect to the whole"), Homeomerous (parts are similar/dissimilar in nature to each other and to the whole to which they belong) and Separable ('"parts can/cannot be physically disconnected, in principle, from the whole to which they are connected"'). Features on the right in (TABLE 1) stand for the author's characterisation of each type in terms of these so-called 'relational elements'.

3 Moreover, they hypothesize that there is a hierarchical ordering:

CLASS INCLUSION > MEREROLOGICAL INCLUSION > SPATIAL INCLUSION

which permits syllogistic transitivity in presence of two of those inclusion types if and only if the conclusion belongs the lower type, as in:

a. Wings are parts of birds (Meronymy)
b. Birds are creatures (Class)
c. Wings are parts of creatures (Meronymy)
d. *Wings are creatures (Class).
So, with respect to the words we are focusing on, they fall into the Portion-Mass class, thus characterised as non-functional, separable and homeomerous to the whole. The latter feature differences Portion-Mass from components and members of collections, as long as a slice of a pie is "pie" and is similar to each other slice and to the whole pie, whilst a window is not like a house nor like another part of it, such as a roof, and a tree is not like a forest nor is it "forest".

They also claim both that portions of masses have arbitrary boundaries and which we can divide and apportion masses by means of standard measures such as inches, gallons, hours and so forth.

2.1.2 CHAFFIN et al. 1988. The authors explode the idea first set in [WINSTON et al. 1987] that relations are abstract concepts composed of relation elements. Relation elements account for relations' characteristics and for the speaker's abilities to make judgements about them; the more elements two relations share, the more similar those relations are.

Basing on psycholinguistic experiments (their own work, [CHAFFIN & HERRMANN 1984]) they provide a classification of 31 semantic relations taxonomized attending to their grade of similarity, where the Part-Whole relation is a family of seven:

(TABLE 2)

| Functional Object (tree-leaf) | PR,PO,A,C |
| Collection (forest-tree)     | PR,PO,H   |
| Group (choir-singer)         | PR,PO,H,S |
| Ingredient (pizza-cheese)    | PR,PO,I,C |
| Functional Location (kitchen-stove) | PR,PO,A,C |
| Organisation (army-corps)    | PR,PO,A,C,S |
| Measure (hour-minute)        | H         |

Features on the right are the relation elements, derived by [STASIO et al. 1985] from linguistic and psychological literature, which, the authors hypothesise, differentiate types of Part-Whole relations among them (there are also other relation elements, omitted here, which distinguish Part-Whole from other families of relations); These elements are defined:

(TABLE 3)

| PRoperty (w2 is a property of w1) |
| POsession (w1 possesses a property or attribute) |
| Attachment (w2 is attached to w1) |
| Component (w2 is a component of w1) |
| Homogeneous (w2's are interchangeable) |
| Social (w2 is part of w1 by social agreement) |
| Locative Inclusion (w2 is "in" w1) |

They further supply experiments which, it is claimed, demonstrate the mental actuality of those elements.

They further give a more detailed sorting of Part-Whole and similar relations, extracted from [CHAFFIN et al. 1988]; this time, decomposition on relation elements is not provided.

---

4 But one may notice that Functional Object and Functional Location share the same set of elements, so there are no distinction between them at this level.
The sorting consists of 31 relations, grouped under Part-Whole, Stuff, Class Inclusion, Representation, Stages and Attributes. Part-Whole itself is divided into Integral Object-Component (6 relations), Event-Feature (2), Object-Topological Part (1), Collection-member (3), Area-Place/Time (3) and Mass-Portion (3). Examples which affect our interests are:

(TABLE 4)

OBJECT-TOPOLOGICAL PART (room-corner, box-side, mountain-foot)
MASS-PORTION: Measure-Unit (mile-yard)
Mass-Natural Tiny Piece (salt-grain, snow-flake)
Mass-Measured Portion (pie-slice, land-parcel)

The proximity clustering provided shows that, attending to the perception of the subjects of the experiment, the strongest distinction among Part-Whole relations is that between portions of mass relations (Mass-Natural Tiny Piece and Mass-Measured Portion) and the rest. The authors impute this to the fact that these relations are homeomerous, and the rest are not.

A clustering of 31 part-terms occurring in the frame "A is a [part-term] of B" due to [TRANSUE 1982] is also provided. It results in the following five major clusters:

(TABLE 5)

(i)- STAGES OF PROCESSES: phase, stage
(ii)- PARTS OF AN EVENT/SCRIPT\(^5\): role
(iii)- PARTS OF PLACES: area, region, zone, section, division
(iv)- PORTIONS OF MASSES (HOMEOMEROUS)
   (iv.I)- BASED ON MEASUREMENT: share, percentage, portion
   (iv.II)- DETERMINED BY THE NATURE OF THE MATERIAL OR OBJECT:
          slice, piece, bit, sliver, drop, grain, lump, scrap, chip, fragment,
          fraction
   (iv.III)- segment
(v)- NON-HOMEOMEROUS PARTS
   (v.I)- OBJECT-COMPONENT: component, feature, part
   (v.II)- COLLECTION-MEMBER: item, member
   (v.III)- STATE-CAUSE: factor, element
   (v.IV)- ingredient

The ability of subjects to sort part-terms make the authors suggest that its meaning, like the relations to which they refer, can be decomposed into more basic elements.

2.1.3 IRIS et al. 1988. On the basis of dictionary data analysis, the authors come to support [WINSTON et al. 1987]'s claim that PART-WHOLE is not a single relation but a collection of them. They find four different types of meronymy:

(TABLE 6)

\(^5\) The authors use here [SCHANK & ABELSON 77]'s notions of 'script' ans 'role'.

6
and conclude that transitivity operates in (b) and (d) but not in (a) and (c).

We can notice two main differences with [WINSTON et al. 1987]'s approach: (i) [IRIS et al. 1988] consider IS-A ("set inclusion") a kind of PART-WHOLE relation while [WINSTON et al. 1987] do not; and (ii) [WINSTON et al. 1987] claim for transitivity along every particular type of meronymy while [IRIS et al. 1988] argue that only Segment-Whole is transitive.

They also straightforwardly reject [CHAFFIN et al. 1988]'s Relation Element Theory arguing biases caused by the ad hoc nature of the original list of examples used with their subjects; consequently, the resulting clustering which in turn is used to create the fundamental relations is biased as well. In their words, "the results obtained by Chaffin and his co-workers indicate how they particular subjects performed on a set of words proposed by the experimenters; it is not clear how this performance relates to native speakers' abilities to identify and label parts and wholes. Nor is it necessary to suppose that in natural situations speakers use Chaffin's 'decompositional strategy' in their decisions".

Instead, they hypothesise that subjects have models or schemata of at least the four kinds of Part-Whole relations listed under (TABLE 6) and that these differ in terms of discreteness and formedness (entitativity) of the whole and the part, and in which came first, part or whole.

[IRIS et al. 1988] conception of Segment-Whole relation implies:

(i) The removability of the part (or the divisibility of the whole).

(ii) Some sense of entitativity is attributed to the part. In some cases the part has a predetermined size and shape; e.g. a slice of bread connotes a particular shape, size and thickness, but in contrast words such as "fragment" do not carry an implication of a particular shape or size.

(iii) The whole has to precede the part (one cannot have a piece of pie before the pie exists).

(iv) The whole may be either discrete or continuous, which is reflected in language by the mass-count distinction (many beans vs. much rice).

(v) Generic terms of mass individuation may extend to abstract, immaterial mass-words (a piece of news or advice) -observation due to [JEPPESEN 33].

(vi) General collective nouns (a pile of sand) or measurement nouns (a cup of rice) are considered to give boundaries to masses -which previously are amorphous wholes. In a more general sense, boundaries can be drawn around segments of a whole.

(vii) An Euclidean geometric conception of space is implicated in how speakers conceptualise things: children developmentally learn that bikes have fronts, forests have edges, etc.

2.1.4 JACKENDOFF 1991. As it is widely known, in Ray Jackendoff's work (Conceptual Semantics) decomposability of concepts is taken for granted. Roughly, both
syntactic structures and lexical items are understood as mental constructs (in the sense of [CHOMSKY 1986]'s I-Language, cf. [JACKENDOFF 1989]) encoded in terms of a finite set of primitives and a finite set of principles of combination generatively configuring the Syntax of Thought.

In the article here sketched he develops a family of conceptual features and functions addressed to globally account for apparently distinct phenomena such as plurality, mass-count distinction, several part-whole entailments, 'Universal Packaging' ("three coffees"), boundary words ("end", "edge", "crust") or the analysis of [Vendler 1957]'s classes of events.

His approach implies a view of a Lexical Conceptual Structure (LCS) of words constructed in terms of more primitive conceptual elements such as Boundedness, Internal Structure, perceptual Dimensionality (following and extending [MARR 1982]'s theory of encoding of object shapes) or Directionality.

An object or event will be +/-B(ounded) according to the claimed speakers' conceptualisation of it as having or not boundaries. This does not entail that an -B entity is absolutely unbounded in space or time but that a speaker uses -B constituent to refer to an entity whose boundaries are not in view or not of concern. Individuals, Groups, Accomplishments or Achievements ("a pig", "a committee", "eat an apple", "to arrive") are +B; while Plurals, Substances, Aggregates or Processes ("pigs", "water", "cattle", "to swim") are -B. Furthermore, something is +/-I (i.e. has has not Internal Structure) depending on whether it is not conceptualised as comprising a multiplicity of distinguishable individuals. Thus, Groups, Aggregates, Plurals or Iterative Processes ("a committee", "cattle", "pigs", "the light flashed until dawn") are +I; while Individuals, Substances or Achievements ("a pig", "water", "to arrive") are -I. Consequently, in [JACKENDOFF 1991]'s terms, Gridding is characterised as a function mapping an +B-I item ("lamb'/animal") into an -B-I one ("lamb'/meat") and Plural, a function mapping +B-I words ("lamb") into -B+I ("lambs").

As long as it somehow concerns to Part-Whole entailments, he posits functions which map:

(a)- masses into +B-I elements, as "a grain of rice" or "a drop of water";
(b)- components into +B+I wholes, as "house" (from "bricks" or "wood");
(c)- wholes into +B-I parts, as "leg" (from "table");
(d)- ingredients into -B-I wholes, as "stew" (from "beef").

Dimensionality is conceived as that speakers conceptualise objects (and also events) as n-Dimensional, i.e. roads or rivers are mentally schematised as lines (1D), carpets as surfaces (2D) or oranges as volumes (3D); a bounded object can also be idealised as a point (0D) -which is the principle that allows cities to be represented by points on maps. Similarly, Accomplishments are 1D while Achievements are 0D. Moreover, objects show a secondary dimensionality, that of their cross-section, so two 1D objects, a ribbon and a tube, differ in that their secondary dimension is, respectively, 1D (a line) and 2D (a circumference). This has implications on words of boundaries, since for instance an "end" may be a 0D part of an 1D item (no matter if we talk either of "the end of a line" or of "the end of a speech"), or even a 1D part of a 2D one ("putting the cup on the end of the table"); this is to say: an "end" is an n-1D part of something nD-bounding its linear axis. The "end" actually may expand a small amount along the axis (namely, it is obviously not always strictly 0D), since "the end of the talk" or "the end of the tube" is not just their not-dimensional boundary but an amount of time or material small enough to be mentally construed as their bounding part.

---

6 More precisely, the author's proposal is the LCS of the 'meat' sense to bear both -B-I features and a function GR(inding) whose argument is the LCS of the 'animal' sense.
For Jackendoff the primitive elements implied in this computing are not given raise by experience; on the contrary, they are innate and they are the stuff by whose means we experience the world.

2.2 COMPUTATIONAL LEXICONS

2.2.1 EDR and WordNet. As far as we know existing large computational lexical data bases treat meronymmycal relations as unanalysed, assumed as uncontroversial, links between words. They just differ in which relations -if any- are included.

    For instance, EDR's Concept Database only includes, among 27 concept relation labels, the merological relation "part of", defined as "whole-part relation; eg. c#wing --> c#bird" [UCHIDA 1990].

    WordNet makes use of three relations for nouns based on those defined by [WINSTON et al. 1987]:

        (TABLE 7)

      HAS_MEMBER/MEMBER_OF (tree-forest, England-Englishman)
      PART_OF/HAS_PART (branch-tree, pint-gallon, England-Manchester)
      HAS_SUBSTANCE/SUBSTANCE_OF (tree-wood)

Transitivity is considered problematic along these relations ([MILLER et al. 1990]), although it is implemented when considered plausible; for instance "cup" is PART_OF "pint", and by inheritance PART_OF "quart-gallon"; "bread" is PART_OF "sandwich", and by inheritance of "meal" as long as "sandwich" is PART_OF "meal".

    As it can be noticed, portion-mass relation (slice-cake) is not included in WordNet.

2.2.2 ACQUILEX LKB (VOSSEN & COPESTAKE 1994). Unlike the databases mentioned above, the Acquilex-II Project provide an analytic approach to some types of meronomy. This approach follows from two of its fundamental idiosyncrasies: (i) Lexical entries are complex featured objects bearing rich linguistic information; and (ii) this information is to be recovered as long as possible from definitions of the words existing in conventional (although Machine-Readable) dictionaries.

    LKB entries bear basically Morphological, Syntactic, Formal-Semantic and Lexical-Semantic features; operations described in [VOSSEN & COPESTAKE 1994] involving

---

7 EDR: the mega-large Japanese-English electronic dictionary project, promoted by the Japan Electronic Dictionary Research, which includes for each language General, Terminology, Concept, Coocurrence and Bilingual dictionaries as well as Corpora; their concept Dictionary, for instance, contains 400,000 words.
8 WordNet: English lexical relational data base containing about 100,000 words organized in about 70,000 word meanings, or sets of quasi-synonyms, based on psycholinguistic grounds.
9 A notation or its reverse is used depending on directionality, namely which word is being charakterised, e.g. 'tree —MEMBER_OF—> forest' or 'forest —HAS_MEMBER—> tree'.
10 The basic purpose of Acquilex-II Project (ESPRIT BRA 7315) is building a multilingual Lexical Knowledge Base (LKB) - whose lexical entries are represented as typed feature structures hierarchically ordered along an IS_A relation to allow inheritance of properties- by means of semi-automatically recovering information from Machine Readable Dictionaries.
meronymy concern the latter. The lexical semantics adopted by Acquilex-II is mainly that of [PUSTEJOVSKY 1991], which posits nominal lexical items to bear a so-called Qualia Structure encoding the four following aspects of their referent: Telic Role (the purpose of an entity, e.g. bread-eating), Agentive Role (its origin or way in which it was created, e.g. cake-bake), Constituency (its component parts) and Formal Role (its physical form).

When analysing dictionary definitions, the authors found that some of them bear complex kernels as in those in (2) which impede or make difficult the standard way of given a word filling it with information by recovering it mainly from the genus of its definition -that is, stating an IS_A link allowing inheritance.

(2)  
a. root 1 : the part of a plant that...  
b. meal 1 : an amount of food eaten at one time.

Taking into account an inventory of words that frequently occur in this kind of structure and do not frequently occur without such of-complementation, [VOSSEN & COPISTAKE 1994] detect (in dictionary definitions) four different classes of relations, one of which is TYPE/KIND (regarded as a version of ordinary IS_A relation) while the other three are merological:

<table>
<thead>
<tr>
<th>TABLE 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUANTITY/MASS</td>
</tr>
<tr>
<td>meal 1 : an amount of food eaten at one time</td>
</tr>
<tr>
<td>band 1 : a thin flat narrow piece of material</td>
</tr>
<tr>
<td>waste 1 : an unused or useless stretch of land</td>
</tr>
<tr>
<td>MEMBER/GROUP</td>
</tr>
<tr>
<td>policeman : a member of a police force</td>
</tr>
<tr>
<td>band 2 : a group of musicians</td>
</tr>
<tr>
<td>COMPONENT/WHOLE</td>
</tr>
<tr>
<td>cockpit : the part of a plane or racing car...</td>
</tr>
<tr>
<td>dolmen : a group of upright stones...</td>
</tr>
</tbody>
</table>

IS_A is the only relation between words (senses) explicitly admitted in the LKB\textsuperscript{11}; the authors refuse to include new links such as HAS_MEMBER and so, since "semantics for these proposed links is not provided" and, they suggest, adopting them would introduce undesirable complications in their task. Otherwise, they look for providing a linguistically motivated account of complex kernels of definitions and infer from this the effects on the construction of LKB entries.

When focusing on QUANTITY/MASS kernels, they define them as "producing an individual from a mass denoting genus term by describing the way in which the entity is to be individuated". The effect will be that the entry so defined, (i) would belong to the type expressing individuals (lex-count-noun) instead of that expressing mass (lex-uncount-noun); and (ii) would inherit all the Qualia Structure of the genus\textsuperscript{12} except for its FORM feature, which will be specified by the syntactic portion-denoting kernel. Thus the entry for meal\textsubscript{1} deduced from its definition (2b) would be:

\textsuperscript{11} Although, obviously, other features of the structure having senses as values are nothing but relations between words, e.g.:

meal\textsubscript{1} <QUALIA:TELIC> = eat\textsubscript{1}

\textsuperscript{12} Structures' inheritance is noted by '<'; see (3).
For the cases in which the mass and the individual readings are senses of the same headword, such as whisky_1 (the liquor, mass) and whisky_2 (a portion, a glass-ful of it), a sense-extending lexical rule 13 is posed which is equivalent to the inheritance treatment given in (3); consequently, the extended sense of whisky will bear identical feature structure as the basic sense except for their type and FORM-feature, which will shift, respectively, from 'lex-uncount-noun' and 'mass' to 'lex-count-noun' and 'portion'.

SYNTAX: ARE PORTION NOUNS NOUNS AS THE REST?

3. [VOSSEN & COPESTAKE 1994]'s approach implicitly gives raise to a key aspect of words of portion which is missed in the other approaches summarised: their necessary syntagmatic projection. Unlike most of the part denoting nouns belonging to other types of meronymy, portion nouns largely underspecify which kind of whole they are parts of; thus they (always) need genitive complementation to fully specify what we are talking about (but see footnote 1).

[Winston et al. 1987], [Chaffin et al. 1987], and [IRIS et al. 1988] don't face this aspect undoubtedly due to the strong lexical-relational nature of their approaches. [JACKEENDOFF 1991] do actually deals in depth with the syntagmatic effects of the word "end", but his work is not concerned with the need of syntactic complementation of portion nouns as a class either. In fact what all of them are concerned 14 is with the psychological entitativity of the elements giving raise to merological concepts. We don't mean to make any claim, at least here, about the psychological status of the elements describing portion words 15. Otherwise, our aim is looking for a linguistic account of their characteristics in order to both provide a description of their lexical meaning and predict their syntagmatic semantic effects.

With respect to this, we think the relational approach of Chaffin and his colleagues is not specially useful, in the sense that dealing with portion terms by putting them together with fully referential nouns and setting there 'PORTION-OF' (or similar) links does not accurately account for important differences between both types of nouns. Portion denoting nouns indeed are nouns from a categorial point of view, but unlike fully referential nouns, they seem both to structurally need for and semantically select an of-complement (in Spanish, a PP headed by the proposition 'de') 16 and, contrarily to other cases of PP complementation, in this case the semantic head of the syntagm seems not to be the structural head, but the complement. At first sight, one could attribute the structure (4b-fig.1) to (4a):

(4) a. Un trozo de naranja
   (a portion of orange)

---

13 Lexical Rules are feature structures representing the relationship between a pair of lexical entries: the output entry is the result of the application of the Lexical Rule on the input one.
14 Not so strongly in [IRIS et al. 88].
15 Although is our belief that linguistic description must necessarily map somehow the speakers' mental processing.
16 At least most of them. Certainly some nouns which seem to naturally fall inside the group of portion-denoting nouns, e.g. andrajo, buncal or carambano, are much more eager to appear without complement than bocado, cala or cara. We will discuss this below.
b. [NP[SPEC un[N trozo[COMPLEMENT/PP de naranja]]]]

```
NP
  \   / SPEC
   /   |
  N'    \
  \   /    \
   N   COMPL/PP
  \  /  |
   trozo de naranja
```

Figure 1.- structure (4b)

but consider (5):

(5)  

a. comeré una naranja  
(FUTURE-eat an orange)

b. comeré un trozo de naranja  
(FUTURE-eat a portion of orange)

c. * comeré un trozo de piedra  
(FUTURE-eat a portion of stone)

The verb ("to eat") semantically selects an edible entity as its internal argument. In (5a) this constraint is satisfied as naranja, the head of the verbal complement, refers to an edible object. But (5b) is also a grammatical sentence; then if we supposed (4b) was the structure for the NP we would be forced to assume that its head, trozo also refers to an edible entity -which is not the case, as we can infer from the unplausibility of (5c). Then we have to assume that the heads of the NPs in (5b) and (5c) must be naranja and piedra, not trozo.

Consequently trozo must be some kind of operator on the head of the NP, not the head itself; hence (4b) is not a correct structure for (4a). It seems more plausible to think of an NP as (4a) as (roughly) bearing a structure as (6):

(6)  

[NP[ADJUNCT un trozo de[NP naranja]]]

17[Hernanz & Brucart 87] suggest this kind of structure (a) for representing similar phrases (b):

(a) [Q algunos de[NP [DET los[N muchachos]]]]

(b) Algunos de los muchachos  
(some of the boys)

but they finally discard the solution in part by arguing that the structure ought to be generalized to handle partitive NPs such as 'a big portion of', 'a meter of' or 'a glass of', which should have to be considered as quantifiers, thus breaking the general principles which govern the structure of NPs in Spanish. Alternatively, they argue for (c):

(c)  

[DET algunos [N[N e][COMPL/PP de los muchachos]]]

assuming the existence of an empty head of the NP coindexed with the head of the complement.
This makes us suggest that portion nouns should not be treated as current nouns but better as
some kind of operators on nouns. With respect to this, the strong parallelism of quantified
sentences in (7a,b,c) to (5b) -here repeated as (7d)-, suggests that portion-denoting
constructions as un trozo de might be considered as (a peculiar sort of) quantifiers since from a
semantic point of view all verbal complements in (7) can easily be interpreted as the quantity of
orange/s to be eaten by the agent, namely three of them, a lot of them, the exact partition half of
one, or an undeterminedly measured partition of one.

(7)  
    a. comeré tres naranjas  
        (FUTURE-eat three oranges) 
    b. comeré muchas naranjas  
        (FUTURE-eat a-lot-of oranges) 
    c. comeré media naranja  
        (FUTURE-eat half orange) 
    d. comeré un trozo de naranja  
        (FUTURE-eat a portion of orange) 

This would lead to posit a structure as (8):

(8) \[ NP[Q \text{ un trozo de} [NP \text{naranja}]] \]

An alternative possibility which would respect for sentences as (d) general principles governing the
structure of NPs in Spanish -keeping as [HERNANZ & BRUCART 87] within a Generative Grammar
framework- would be assuming a standard Surface Structure as (e) and positting restructuring movement of
constituents to (f) in the Logical Form:

(0) \text{ un trozo de naranja} 
(0e) \[ NP \text{ un trozo de naranja} \] 
(0f) \[ NP[ADJ/QP \text{ un trozo de} [NP \text{naranja}]] \]
Another interesting possibility for a solution is that suggested by [CHOMSKY 1981], (cf. [HERNANZ & BRUCART 1987]), where in order to reflect the lack of semantic concreteness of the preposition 'of' (parallel in Spanish: 'de'), he posits consider it as an expletive case marker. The elision of 'de' in colloquial Spanish phrases as those in (9) seems to support this.

(9) a. un trozo naranja
    (a portion orange)
    [a portion of orange]

b. un cacho pan
    (a portion bread)
    [a portion of bread]

On this view portion terms could be characterised as nouns heading adjunct (partitive) quantifier-NPs governing NPs in genitive (normatively) marked by preposition 'de'/'of'; roughly:

(10) \[ \text{NP}[Q\text{NP un trozo [CASE-MARKER de[NP naranja]]}] \]

![Figure 3.- Structure (8)](image)

![Figure 4.- Structure (10)](image)

3.1. Relational approaches don't suit for portions. We leave here open the issue of the structural representation of such phrases; but all this discussion above stands to say that, contra relational approaches, whilst on the one hand linking 'tree' to 'forest' or 'musician' to 'band' by means of a conceptual link such as 'MEMBER-OF' seems to be a productive linguistic generalisation inside the class of nouns, on the other hand linking 'slice' to 'bread', 'pie', 'fruit' or any other 'sliceable' object by means of a 'PORTION-OF' relation looks much the same as stating a cross-categorial relationship. Linking 'grain' to 'salt' or 'slice' to 'pie' by means of some sort of PART-WHOLE relation is much as linking 'stones' to 'pile' by some MEMBER-GROUP relation; it is pretty unprecise: properly speaking neither 'grain' is a part of 'salt' nor 'pile' is a group of stones, but 'a grain of salt' is a part of 'salt' and 'a pile of stones'
is a group of stones (see fig. 5). Therefore these relations are not to be encoded in a lexicon, as long as they are not relations between lexical units, but between lexical units and syntagms.

(a) pie

\[ \text{PART-WHOLE} \]

slice

(b) pie

\[ \text{PART-WHOLE} \]

a slice of pie

figure 5

We suggest instead that portion nouns must be dealt with in a lexicon by

(i) considering them a class which is distinct of that of the overall nouns,

(ii) describing them internally according to their intrinsic meaning components (form denoted, nouns selected and/or expressed by default, and so forth); and

(iii) leaving to syntagmatic rules the task of handling this information to construct the new concept resulting of the composition of the portion-noun along with the full-entity-noun which complements it.

With respect to point (i) our belief is that the class of portion nouns, as a syntactic approach:

(a) show strong similarities with (partitive) quantifiers inasmuch they are not heads of argument NPs but operators which place a measure or quantity on actual heads -but instead usual quantifiers they do not only specify measure but also other pieces of meaning, e.g. form-;

(b) nevertheless they are not categorially quantifiers but (quantifier-making) nouns: they are heads of NP structures which in their turn are the actual (partitive) quantifiers of the lower level NPs;

(c) they govern genitive structures which are case-marked in Spanish by the preposition 'de' and in English by 'of' -but despite this, some of them neither seem to overtly select heads at all nor govern genitive complements; instead they seem to have lexicalized the semantic head thus expressing not only form or quantity but also the stuff the portion is made of.\(^{18}\)

---

\(^{18}\) There is about this issue the interesting example in Spanish of the term 'quinto' -which we think to be parallel to the English term 'pint'. This noun denotes a specific quantity of beer usually draught in bars, namely 1/5 L. We found it interesting as we are able to trace its process of lexicalization of the stuff it refers to. Originally 'un quinto' or 'una quinta (parte)' are partitive terms denoting the 1/5 part of anything. In our decimal culture we can suppose the term was steadily specialised to denote 1/5 of a liter of any liquid. Somewhere in the time people in Spain progressively gave up drinking wine in bars in favour of the trend of drinking beer and, a standard measure of beer to be draught has come to be 1/5 of a liter. The economy of language has rapidly discarded the term 'un quinto de litro de cerveza' (1/5 L. of beer) in favour of 'un quinto de cerveza' (1/5 of beer) -lexicalizing, or tacitly making understand, the measure unit 'one litter' inside 'un quinto' -; in a second step, the convention made 'un quinto' to even lexicalize the liquid of reference. The result is that, nowadays, if you sit in some bar in Barcelona and order 'un quinto', you'll be served a bottle of 1/5 of a liter of lager.
SOME REMARKS ON [VOSSEN & COPESTAKE 1994]'S APPROACH.

4. From the previous section must be inferred that portion nouns should not be dealt with relationally but grammatically. Consequently our approach is implicitly coincident with that of [VOSSEN & COPESTAKE 1994] in the sense that what is really needed is an account for the semantics of portion nouns as operators '-relators' in their terms. Then now we shall discuss that approach and set some remarks on it.

First we have to notice that it seems obvious that nouns denoting portions must encode in their meanings something more than "this is a portion of X", otherwise the display of terms (LIST 1) is a subset of would be largely uneconomic, therefore language-speakers- would not allow them to pervade time -other authors here discussed, specially the seminal work of [WINSTON et al. 1987], had already shown that these words of course also encode something else than "this is a part of X". Once this point of view assumed, we think the effects predicted by [VOSSEN & COPESTAKE 1994] of portion words19 as causing a shift on the FORM slot of the output to the type "portion" is too coarse-grained a generalisation. As they suggest, further parsing of -in their framework- the rest of the dictionary definition (the Aristotelian 'differential') is bound to provide the rest of distinctive semantic features; for instance in:

(11)  band 1 : a thin flat narrow piece of material.

the word "piece" is entailing the shift of the FORM type from that previously assigned to the genus ("material") to "portion", whilst adjectives "thin", "flat" and "narrow" would provide further semantic details.

However, consider:

(12)  a. waste 1 : an unused or useless stretch of land

b. lámina 1 : plancha delgada, esp. de un metal
(slab thin , esp. of a metal)
(thin slab, esp. of a metal)

c. Pásame una rebanada del de centeno
(Give-me a long-thin-slice of-the of rye)
(Give me a slice of the rye one) {understood:bread}

If our non-native knowledge of English doesn't trait us, regarding the word "stretch" in (12a) as denoting just a "portion" of land -further modified by "unused or useless" to build up "waste 1"-, doesn't grasp the idea that such portion of land can not have an extension of, for instance, five square inches -it should be remarkably larger. In the same way, neither plancha in the dictionary definition of "lámina 1" in (12b) can be regarded as a spherical portion of anything, nor rebanada in (12c) may be half a bread. Therefore a semantic specification of such words which be more fine-grained than just telling they are "portions" is needed in order to give a plausible account of those phrases where they appear in -which by the way would be useful to a more effective recovering of information from dictionary definitions.

With respect to the other type shift predicted by [VOSSEN & COPESTAKE 1994], that of the general word-type from "mass" to "individuated", "-lex-uncount-noun" to "lex-count-noun" in LKB terms-, it undoubtedly is a fine strategy for dealing with lexical entries, but else

19 Is not to be forgotten that [VOSSEN & COPESTAKE 94]'s aim is analyzing dictionary definitions to grasp the semantic characteristics of words thus defined. For us it is not relevant here if phrases containing portion words eventually set resulting on a lexical item, as long as while we focus on portion words they focus on words defined in dictionaries by means of them. Notwithstanding the same kind of research underlies both goals since the semantic effects of portion words have to be established no matter if the phrase "a thin flat narrow piece of material" corresponds to an English existing word, or not.
one has to notice that, as in (12c), many portion nouns may as well generally apply on individuated objects; namely, one can either ask for a slice of "bread" or for a slice of some particular piece of bread. The allowance of use of individuating determiners as in (12c) or (13) provides evidence of it.

(13) a. Guardaba un jirón de su vestido de novia 
(PAST-keep a rag of her dress of bride) 
[She kept a rag of her wedding dress]

b. Dame una loncha de ese jamón
(Give-me a long-wide-thin-portion of this ham)

This precision is important since masses are not supposed to bear a definite form but the form of (individuated) portions sometimes depend on the form of the whole, as in [JACKENDOFF 1991]’s examples (14a-c) or other (14d-e):

(14) a. The end of a line

b. The end of a beam

c. The end of a table

d. Una cara de un disco 
[a side of a record]

e. La superficie de una mesa 
[the surface of a table]

f. La punta de la lengua 
[the tip of the tongue]

Another remark we want to make about [VOSSSEN & COPESTAKE 1994]’s approach concerns to inheritance. Words defined by a portion term of-complemented by the genus are regarded as inheriting by default from the genus all their Qualia except for FORM -but of course further properties provided by the ‘diferentiae’ will overwrite the appropriate slots. In more general terms this means that, in absence of further specification, concepts expressed by the pattern

(15) [portion-noun] of [whole-noun]

would inherit the semantics of the concept expressed by the noun of the whole -except for form. This is not always completely true as long as, for instance, although on the one hand there is no doubt a slice of bread keeps being bread, therefore if bread has been brought about by baking then the stuff of which the slice is made of has also been brought about by baking; on the other hand it is not less true that one can also say that the slice itself has been brought about by cutting it off the bread. Consider the following Catalan sentences:

(16) a. En Joan va fer el pa i en Pere va fer les llesques 
(DET Joan PAST-make the bread and DET-Pere PAST-make the slices)

b. En Joan va coure el pa 
(DET Joan PAST-bake the bread)

c. En Pere va tallar el pa 
(DET-Pere PAST-cut the bread)

d. En Joan va tallar el pa 
(DET-Joan PAST-cut the bread)

e. En Pere va coure les llesques 
(DET-Pere PAST-bake the slices)
The assertion (16a) entails (16b) and (16c), but it doesn’t entail (16d) and (16e)\(^\text{20}\).

All this is to say that, keeping inside Pustejovsky’s Theory of Qualia terminology, indeed from one point of view, a word such as “slice” inherit from their whole the Agentive Role “bake 1” or so; but from another point of view “slice” has to bear an Agentive Role such as “cut off” which is not inherited but inherent to it. In other terms, it is not always sure that wholes may be substituted by portions in the same grammatical contexts, e.g., argument positions of some verbs.

**SEMANTICS: ON THE COMPLEX BEHAVIOUR OF PORTION NOUNS.**

5. In 3 we outlined our syntactic approach to portion nouns. Here we are going to sketch their semantic properties. From this point of view portion terms are quite a complex object as they seem to be a sort of cross-roads item -sharing properties of several lexical categories.

[A]. As we posed above, they are quantifiers in the sense that they specify quantity or measure -the only difference is that quantifiers are usually considered to quantify over 1 (two, three, some, many) while portions (at least portions of solids) work on ranges below the unit.

This measure is not always underspecified (e.g. measure:portion), in fact always every portion term place any further constraint on quantity or size (a stretch of land is a relatively big portion of land, a grain of salt is a very small portion of salt -in fact, probably the smaller which is perceivable).

[B]. Inasmuch they are quantifiers, they are specifiers: the portion is attributed entititavity [IRIS et al. 1988] and even when applying on mass nouns the portion turns to be an individuated item -with the syntactic consequence of shifting uncountable nouns to countable [VOSSEN & COPESTAKE 1994].

[C]. Nevertheless portion nouns still are nouns: the portion itself has the semantic feature of being obtained by means of some process of its own, different of that which brought about the whole (see the discussion about the examples on (16)). This is a property of nominals: the Agentive Role [PUSTJEVOSKY 1991].

But they also show an adjectival behaviour: they do not just quantify the noun but also:

[D]. They predicate properties of it, as shape, size or other (a ‘lamina’ is quite thin, a *rebanada* is oval and not thicker than, say 1 cm.).

These and other factors usually conflate in a lexical item in ways which are difficult to discern: a *rebanada de pan* is a quite small portion of a bread, has a determinate form, and it has been obtained by cutting off; a *trozo de pan* might be bigger or smaller than a *rebanada* but certainly it is less than half a bread, and neither has got a determinate form nor places precision on which process of detachment it has been obtained; a *mendrugo de pan* might look like a not very big *trozo de pan* but the bread is stale.

[E]. Being predicatars, they semantically select the class of nouns to be their arguments, namely, they bear selectional restrictions;

[F]. As it happens to adjectives, by use or convention there might occur the lexicalization of the (semantic content of the) noun they are modifying.

\(^{20}\) (16c), by the way, seems to be quite an implausible sentence.
Let us see below in (17) some examples of parallelisms [D,E,F] between adjectives and portion nouns.

(17)  

a. Un listón de madera  
(a lath of wood)  
b. Una madera alargada  
(a wood elongated)  
c. Un pedazo de piedra  
(a piece of stone)  
d. Una piedra pequeña  
(a stone small)  
e. Un mendrugo de pan  
(a hard-crust of bread)  
f. Un trozo de pan viejo  
(a portion of bread stale)  
g. Una rebanada de pan  
(a slice of bread)  
h. *Una rebanada de agua  
(a slice of water)  
i. Cortar pan  
(to-cut bread)  
j. *Cortar agua  
(to-cut water)  
k. Pan duro  
(bread hard)  
l. *Agua dura  
(water hard)  
m. Un billete verde  
(a bank-note green)  
n. Un verde  
(a green)  
o. Una jarra de cerveza  
(a mug of beer)  
p. Una jarra  
(a mug)

About [D], Predication of Properties: In (a) listón supplies the information that the piece of wood is elongated; alargada in (b) does the same. In (c) pedazo informs that the piece of stone is small; the adjective pequeña in (d) also do that. In (e) mendrugo qualifies the piece of bread as stale, as viejo does in (f).

About [E], Selectional Restrictions: A rebanada is a portion obtained by cutting it off the whole (g,h); one can do that with solids (g,i) but not with liquids (h,j); the same thing happens with
the adjective duro: solids can be duro, liquids can not -except for the specialised chemical term agua dura, said of water containing Ca and Mg ions, a technical detail which falls out of the scope of general linguistic accounts.

About [F], Lexicalisation of the argument: in Spain the green note of 1000 pts. has become by use and convention a verde(m,n); the same way a mug of beer is everywhere known as and asked for a jarra (o,p); in both cases the structural and semantic head, the noun, has been conflated into the predicator, showing that portion terms, as adjectives do, can be nominalised.

GENERALISATIONS: OUTLINING A TAXONOMY OF PORTION TERMS

6. Considering the most salient pieces of meaning discussed up to here, we find five major classes of portion nouns: Container, Pre-Segmented, Boundaries, Detached and Modelled.

   Roughly, we can think of them as follows:

   (i) Container portions are those which are held in a container, thus they bear no shape of their own, the rest of types of portions do;

   Detached and Modelled portions did not exist in any way before they were separated of the whole, Pre-Segmented and Boundary portions did, hence we will call the former "Created Portions" and the later "Pre-Existent Portions";

   (ii) Detached Portions have been brought about by a process of detachment:

   (iii) Modelled Portions have been brought about and given form by some process of modelling a mass:

   (iv) Pre-Segmented portions either may have been already detached from the whole or may still be attached to it -but anyway the portions are units of a naturally pre-existing internal structure of the whole;

   (v) Boundary portions are elaborations of natural boundaries of the whole and are conceptualised as those boundaries [JACKENDOFF 1991]; as well, they may still be attached to the whole or not.

     Pre-Segmented portions are basic-level concepts -this is, they are conceptualised as a whole, bearing complex unschematisable forms; the rest are mostly conceptualised image-schematically -this is, as elaborations of pre-conceptual images such as container or periphery [LAKOFF 1987] or of basic geometrical concepts such as line, circle and so on [IRIS et al. 1988].

_____________________________
21 The distinction between basic-level and image-schematic concepts, as well of the notion of global conceptualisation -gestalt- are corner stones of Cognitive Semantics, cf. [LAKOFF 87].
Let us now examine each class in some detail. Certainly you will notice that some examples fit better than other in the class, even that some seem to share features of more of one class. The approach should be taken as a generalisation submitted to further analysis.

(i) CONTAINER PORTIONS
   Examples: *cucharada*, *jarra*, *puñado*, *taza*, *vaso*.
   Features: They select uncountable nouns and plurals -[JACKENDOFF 1991]'s [-B] nouns. They are methonymycal sense-extensions of nouns of containers. They denote no specific form unless one considers as their form that of the container. They clearly specify quantity: that which the container can hold. It may be considered as not bearing Agentive Role as they have not been obtained by means of a particular action; from another point of view one can consider they have been obtained by means of filling the container.

(ii) DETACHED PORTIONS
   Examples: *barra*, *cacho*, *calo*, *escalope*, *filete*, *fragmento*, *jira*, *jirón*, *lasca*, *lengua*, *listón*, *loncha*, *lonja*, *mendrugó*, *orejón*, *parcela*, *pedazo*, *pizca*, *raja*, *rebanada*, *región*, *rodaja*, *redondo*, *sección*, *segmento*, *solomillo*, *taco*, *tira*, *trozo*, *zurullo*.
   Features: They select solids, either masses or individuals. They bear image-schematizable form (shape and size). They somehow constrain a range of quantity relative to the whole. They bear an Agentive Role which is some process of detachment.

   Nouns denoting precise partitions, as *cuarto*, *mitad* or *tercio* are an homogeneous sub-cluster of Detached Portions being characteristic because quantity is conceptualised by reference to exact arithmetic partitions.

(iii) MODELLED PORTIONS
   Examples: *carámbano*, *lámina*, *lingote*, *oblea*, *placa*, *terrón*.
   Features: They select solid masses. They also bear image-schematizable form (shape and size). They bear an Agentive Role which is some process of modelling of a mass (e.g., a

---

22 For instance, the terms carámbano, lágrima and gota seem not to fall clearly inside any of these classes. They could be considered as Detached but it is problematic considere them as obtained by some process of detachment (Agentive Role). Moreover, lágrima and gota could be considered Pre-Segmentied, as all liquids may be thougth as composed of minimal perceivable elements called drops; nonetheless this characterisation seems not clear for us.
terrón de azúcar [a lump of sugar] has been created and given its orthoedrical form by taking sugar and modelling it).

(iv) PRE-SEGMENTED PORTIONS
Examples: brízna, gajo, grano, hoja, mechón, migaja.
Features: They select masses or individuals\(^{23}\), which hold a pre-existent internal structure -[JACKENDOFF 1991]'s [+I] nouns. Both their form and quantity settings on the whole are basic-level conceptualised. It is doubtful to attribute them an Agentive Role, as they can remain attached to the whole; of course they may have been detached from it, but one can not get this knowledge only from the lexical unit-discourse knowledge is needed.

(v) BOUNDARIES
Examples: base, capa, cara, cima, corteza, cuscurro, esquina, exterior, extremo, morro, película, perfil, pie, punta, superficie.
Features: They mostly select individuated solids -but one can also talk about the superficie [surface] of the sea. They are conceptualised as minimal elaborations of boundaries -one of their dimensions is perceived as close to non-existence. They also seem not to naturally bear an Agentive Role.

FEATURES FOR THESE GENERAL TYPES.

7. Although the taxonomy here presented is not grounded in structure inheritance but in basic cognitive distinctions for these type of concepts, we hypothesise that cognitive distinctions will map to a hierarchy of feature structures ordered by subsumption. In (Figure 7) a featured taxonomy of portion terms is sketched. As ordered by subsumption, features are mentioned for each node either when they are introduced at that level or when they acquire a range of values which is more specific than that born by its mother node -otherwise it is to be understood that features and type of values are inherited from the mother node.

7.1 ABOUT TYPES. We assume that both Containers do not specify form and Pre-Existent do not bear an Agentive Role of their own as these features are considered not linguistically relevant for these sorts of terms\(^{24}\).

For the Pre-Existent node (that which subsumes Pre-segmented and Boundary portions) we extend the notion of Basic-Level Concept -which is used in Cognitive Grammar to express Form- to the feature Quantity, in the sense that, for this type of portions, quantity seems not to be calculable as a partition of the whole but instead globally conceived; and also it is straightforwardly related to the form of the portion.

This way, for Pre-Segmented portions, the quantity or measure expressed by e.g. a gajo of a fruit corresponds to its shape and size (Form). For Boundary portions, the quantity which these terms express is, by their own nature of being a boundary, a minimal quantity of the whole -we can not properly say that the measure of an "end" is 1/32 or 1/64 of the whole, but instead, the "end" of something is a minimal portion of it (e).

\(^{23}\) They mainly select solids, but we can consider gota [drop] as belonging to this category if a liquid is entertained as previously composed of drops.

\(^{24}\) As it was discussed above, obviously something filling a container may be assumed as both adopting its form and being brought about by the process of filling; moreover a Pre-Existent portion may have been detached from the whole -but also it may still remain attached to it. But in all cases we have not found linguistic consequences of these phenomena, hence, by the moment, we find no reasons for including these data in a lexical-linguistic account like this.
7.2 ABOUT FEATURES

[A] (quantity): A terminological precision is to be set here: we understand Quantity in nouns of portions as a parameter set by relation to the whole, while Size refers to the portion itself and it is embodied into the concept of Form - we understand Form as a description of the Shape and Size of the portion considered alone, with no longer relation to the whole from which it has been taken apart. Examples of different constraints on Quantity are shown in (18): a bread may be divided in two halves (18c), in, roughly a dozen of trozos (18b) and in thousands of migajas (18a). This is the kind of information portion terms carry with them which we call Quantity.

(18)  
a. una migaja de pan  
(a crumb of bread)  
b. un trozo de pan  
(a piece of bread)  
c. medio pan  
(half-a bread)  

Every partitive term carries information about quantity more or less precisely: those deriving from containers (‘cup’, ‘spoonful’) are quite precise, as they refer to the measure already specifically denoted by the container; those directly expressing measures are of course even more precise by their own nature (‘a third of’); those expressing detached portions of solids (‘slices’ etc.) are not extremely precise but they place the Quantity value between some reasonable limits (e.g. much less than a half but more than a little bit); terms conceptualised as
boundaries ('end', 'tip', 'surface') seem not to be much involved with measurement as, being construed as boundaries, seem to tend to zero in some or all of their dimensions.

[D] (agentive): Instead of the claims of full transitivity of properties attributed to the portion-whole relation made by several authors discussed, portions which have been obtained by some kind of detachment bear an Agentive Role associated to that process.

Examples of terms bearing Portion-Agentive Role are displayed in [19]. (19a) have been obtained by cutting with a knife; (19b) by tearing; (19c) by splitting.

(19)  
a. loncha, cala, raja  
b. jiron, jira  
c. fragmento

[D.1] (form): Talking about form, one could a priori think that there are two kinds of portion terms: those which specify form per se (e.g. rebanada: oval; gota: spherical) and those which underspecify it (e.g. trozo, fragmento). But it is only partially true: portion terms not always specify form due to their own lexical characteristics: some of them specify form due to characteristics of the whole they are portions of.

As on the one hand a carámbano is long-pointed or a tira is elongated per se, on the other hand rodajas are circular because they are sections of, say, oranges, and oranges are spherical; rebanadas in Spain are oval because of the usual form of breads, but when the bread has a polyhedral form, rebanadas are square; and gotas or lágrimas are more or less spherical because they are free small portions of liquids and physical properties of liquids make them freely adopt this form in absence of further constraining forces.

Certainly neither linguistics is intended to deal with physics -in fact most of speakers are not physicians- nor the form of very little portions (drops, grains, etc.) seem to be linguistically or perceptually relevant to speakers, thus for these cases it seems sensible to us to let form underspecified -as one would do with perceptually relevant portions as trozo or fragmento which underspecify form--; but it has to be kept in mind that portion nouns might encode form by at least the following means:

(i)- denoting directly form by their own lexical meaning, probably when applying on mass nouns of solids (lámina, loncha, plancha)

(ii)- denoting operations on formed individuated wholes which result in specific forms of their portions (rodaja, rebanada, segmento)

(iii)- applying on masses, one can consider that form is indirectly when they are either metonymies from containers (vaso, copa, jarra) or container-derived terms (cucharada, bocado)

(iv)- denoting no form at all (trozo, fragmento, porción).

Sometimes a portion term may make simultaneous use of (i) and (ii), namely, a rodaja is round because it is a section of a sphere, but its intrinsic lexical characteristics also imply that it is considerably thin with respect to the plane of the section.

[D.2] (other predications): Here are some examples of predication of properties (apart from form and quantity) carried by some portion terms:

(20)  
a. escalepe : the portion of meat has been fried  
b. filete, redondo : they correspond to an specific part of the animal.  
c. mendrugo : the bread is stale.  
d. orejón : the fruit has been dried

[E] (selectional restrictions): It seems an obvious matter, but it has to be noticed that
while some portion terms may apply on a very wide range of entities, other select them in a very precise way. We can see the gradation in (21):

(21)  
a. trozo = every solid  
b. vaso = every non-solid  
c. fragmento = every solid which can be split  
d. lamina, plancha = metals  
e. lasca = stones  
f. cala = fruits  
g. filete = edible animals  
h. gajo = pre-segmented fruits  
i. jira, jirón = clothes  
j. carámbano = ice  
k. mendrugo = bread

[F] (lexicalisation of the whole): We would not take as portions fully whole-lexicalising nouns as andrajol, asqu, añojal or bancal. They appear in the corpus listed in (LIST 1) because, as the list has been initially extracted from the dictionary by looking for words including nouns as 'piece', 'portion' or so in their definitions, these are defined by these means. But these words -although sometimes in the past probably did- may no longer appear in patterns [[PORTION] of [WHOLE]] (vid. 22) as they have both fully lexicalised the whole and lost their partitive meaning. In other words, today they are as nouns as 'dress', 'fire' or 'field'. Thus, as this paper is intended to deal with portion terms as predicates/operators and these are not, we simply do not face this cases and let it apart to be represented as common nouns.

(22)  
a. *Un andrajol de ropa  
(a rag of cloth)

b. *Un asqu de madera  
(a red-hot-coal of wood)

c. *Un añojal/Un bancal de terreno  
(a special-portion-of-land of land)

A similar but different phenomenon is that of portion terms which, as those above, in some cases appear in isolation thus having lexicalised the meaning of a precise kind of whole (23a,d), but -instead of those above- as well may still appear in patterns [[PORTION] of [WHOLE]] (23b,c,e,f) operating on nouns which can be either that ones they usually lexicalise (23b,c) or other of the same type (23c,f).

(23)  
a. Beberé una jarra  
(I-will-drink a jar)  
{understood: of beer}  
{*understood: of milk/wine/water...}

b. Beberé una jarra de cerveza  
(I-will-drink a mug of beer)

c. Beberé una jarra de leche  
(I-will-drink a jug of milk)

d. Comeré un filete  
> (I-will-eat a sirloin)  
{understood: of beef}  
{*understood: of pork/chicken/fish...}

e. Comeré un filete de ternera  
(I-will-eat a sirloin of beef)
f. Comeré un filete de filetán
   (I-will-eat a sirloin of halibut)

We suggest approaching these terms as sense-extending by convention (thus culture-dependant hence non-predictably) from operators/predicators to common nouns. We posit for these terms to bear two parallel lexical entries, one as predicators/operators, another one as nouns. The sense extension responds to the general schema (24) but possibly it has to be attributed case by case as it can not be deterministically predicted.

(24)

syntax: [[NP[N Portion-Term]]] of [X]
selecting: items of type X
semantics: form + quantity + ...

--> 

syntax: [N Portion-Term]
semantics: form + quantity + ... + semantics of one precise item of type X

OUTLINING LEXICAL ENTRIES FOR PORTION TERMS

8. Some entries exemplifying the classes posited in figure 7 are sketched below. The following general details should be noticed:

(i).- Representations here presented are formalism-independent; they intend to reflect the speakers' knowledge of this kind of words which are relevant for linguistic processing.

(ii).- About the syntactic representation, we assume that knowledge about the syntax of a term is not just knowledge about its category, but as well knowledge of its structural frame; namely, we assume that the speaker's knowledge of the word trozo involves knowing which the word appears in contexts as un trozo de X. This syntactic outline is coherent with approaches such as Categorial Grammars [MORRIL 1994], Tree-Adjoining Grammars [JOSHI & SHABES 1991] or psycholinguistic models which give support to connexionist models of Language Processing [MACDONALD et al. 1994]. The tree we use for representing the syntactic structure is that of figure 4 (Section 3 above).

We assume that the knowledge of the structural frame of nouns o portions also involves the knowledge of the semantics of the noun with whom it combines; namely, speakers not only know that the environment of trozo is un trozo de X but also that X is a solid object. This implies that syntactic knowledge (and thus syntactic representation) overlaps knowledge of Selectional Restrictions.

The model of syntactic representation adopted by Acquilex-II is fully compatible with that here outlined, as long as in Acquilex-II Lexical Representation Language the representation of the word involves the representation of its structural frame and also the semantic characterisation of the signs which appear there.

(iii).- Semantic information combines the conceptual calculus of [JACKENDOFF 1991] with the notion of conceptualisation of the form developed in Cognitive Grammar [LAKOFF 1987], which assumes concepts as belonging to two basic categories: Basic-Level and Image-Schematic (as it was pointed above). We also make use of the notion of Agentive Role
developed in [PUSTEJOVSKY 1991]'s Theory of Qualia (see Section 2.2.2).

We believe that all three approaches are compatible and account accurately for a wide range of phenomena which, possibly, each approach could not account for independently.

(iv).- We think as specially useful [JACKENDOFF 1991]'s concept of minimal spatial elaboration (represented by the symbol \( e \)). Jackendoff uses it for describing nouns of boundaries, positing that speakers conceptualise one of the dimensions of a boundary as a minimal portion of the object: that dimension which in ideal geometric terms would be quantified by the value Zero. Reality is not made up of pure geometrical objects: all real objects do have three dimensions, but it is assumed that speakers, when conceptualising, make geometric-like idealisations. This way, sometimes speakers think of objects as having some of their dimensions in a minimal degree, tending to Zero; this is what is expressed by means of the symbol \( e \). Consequently, a sheet is conceptualised as a plane (two dimensions) - the third dimension is perceived as minimal (\( e \)). It does not happen the same with words such as rodaja or rebanada where the thickness of the portion is conceptualised as quite small -but not as perceptually irrelevant. So, for the third dimension of these terms one should give values standing for "very small size" -but not \( e \), as thickness of rodaja or rebanada is certainly small, but perceptually relevant.

(v).- With respect to relevant dimensions, inasmuch that sort of measures seem to be calculable (between fuzzy ranges) from either the size of the rest of dimensions of the portion or the size of the whole, we provisionally pose represent them as fuzzy values or intervals related to those magnitudes.

This notation will be understood with the help of figure 8. In the gradation of a partitive quantity from nothing to all we can place some reference points (for ease of simplicity we will only take 0, 1/2 and 1); we can see (precise) reference points as generating around them (fuzzy) reference regions (0 generates \( e \), 1/2 generates half). Hence, ranges of quantities falling in reference regions will be notated by the name of the region (e.g. "a little bit" will be notated \( e \)), and ranges of quantities falling between reference regions will be notated by intervals of them. These intervals, as in the usual mathematical notation, may be open or closed; open intervals do not include the reference values in the interval, closed intervals do; open intervals are notated by means of a pair of values enclosed in brackets; closed intervals enclose them in square brackets. This way, "more than a little bit but less than a half" will be notated \( (e, \text{half}) \); and "half or more" will be notated \([\text{half,all}]\).

![Figure 8: Fuzzy Quantities](image)

Using this notation we will say that the Quantity of a trozo is \( (e, \text{half}) \) of the whole; and that the thickness of a rodaja is \( (e, \text{half}) \) of the measure of the plane formed by its first two dimensions (if it were less thick it would not be a rodaja but a lámina (sheet); if it were thicker, it would be a trozo) - see figures 9 and 10.

(vi).- It will be noticed that some entries in the examples show a percentual value associated to the type of noun selected by the Portion Noun. It is well known that semantic knowledge of
predicators involves the typical context of use of the word: selectional restrictions and collocations. This combinations are not regular, predicators combine with some (types of) words more usually than with other ones. Several recent approaches in Natural Language Processing (NLP) encode these kind of preferences by means of weights attributed to the different types of words selected by predicators and pose to retrieve this knowledge from corpora by different methods [RIBAS 1994]. Once enriched with this information, NLP systems might make use of it for different tasks involving decisions on meaning, such as ambiguous constituent attachment, sense discrimination or ellipsis solving.

As we have shown, Portion Terms are predicators, thus it will be useful representing this kind of knowledge in their lexical entries -terrón almost always combines with azúcar [sugar]; gajo most of the times with naranja [orange] and sometimes with limón [lemon]; etc.

Notwithstanding, percentages attributed to the examples below are no actual values, but simple guesses which stand there just to be examples of the kind of information we think a lexical entry of Portion Terms should bear.

8.1 EXAMPLES. Examples presented are trozo, rodaja, migaja, terrón, extremo, gajo and jarra. The first three terms fall within the class 'Detached' -migaja not so clearly, see discussion below-; terrón, extremo, gajo and jarra are examples of -respectively- 'Modelled', 'Boundary' 'Pre-Segmented' and 'Container' portions.

**Representation of TROZO**

<table>
<thead>
<tr>
<th>CATEGORY / STRUCTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
</tr>
<tr>
<td>Q/NP</td>
</tr>
<tr>
<td>case marker</td>
</tr>
<tr>
<td>NP</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>&quot;de&quot;</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>&quot;trozo&quot;</td>
</tr>
<tr>
<td>X ← type: solids</td>
</tr>
</tbody>
</table>

**SEMANTICS**

- quantity: (e. half) of X
- form: underspecified
- agentive: detachment

**figure 9.-** lexical entry of TROZO

[fig. 9] trozo. Form is not specified in this word -if it were, the portion could be named differently (tira if it were elongated, rebanada if it were plane-like and cross-cut, etc.). Form
underspecification in *trozo* makes this word subsume other kinds of portion nouns (every *rebanada* is a *trozo* of bread, but not every *trozo* of bread is a *rebanada*).

The Quantity of *trozo*, as it is explained above, is defined by relation to the measure of the whole it is a portion of: X.

It bears Agentive Role as *trozos* can not exist before being detached of the whole. Thus, the semantics of the construction *un trozo de X* would bear the Agentive Role relative to the Portion -pan [bread] would bear “bake” or so as Agentive while *un trozo de pan* [a portion of bread] would bear “detach”.

**Representation of RODAJA**

**CATEGORY / STRUCTURE**

```
NP
  Q/NP  case marker  NP
    |       |       |
    N     "de"     N
\   /\   /\   /\  type: individuated solid  \\
"rodaja" X   \\
             form: spheric, cylindric
```

**SEMANTICS**

- quantity: [ ]
- form 1: DIM 1&2: a cut across the axis of X [ ]
- DIM 3: (o,half) of [ ]
- agentive: cut

**figure 10.** lexical entry of RODAJA

**[fig.10] rodaja.** In order to set Selectional Restrictions, in this word not only it is relevant the semantic type of the whole, but also its Form. Rodajas are round because they are cross-cuts of objects which, when cross-cut, the result is circular. Consequently *rodaja* involves semantic selection of nouns, not only belonging to a specific semantic type but also bearing a Form which is (conceptualised as) spherical or cylindric.

Quantity is co-indexed with Form as long as one can not know how much of an object is a *rodaja* of it. You can cut seven or eight rodajas from an orange, but forty or fifty rodajas of a long sausage. It is precisely the size of the *rodaja* -more specifically its thickness, the 3rd dimension- which hints you "how much of an orange" is a *rodaja* of orange. This is the reason why we pose that in *rodaja* Quantity and Form share their values -because the Quantity of *una rodaja* might not be calculated from the dimensions of the whole, but from those of the portion itself.

**[fig 11] terrón.** This is one of the cases in which Form is directly schematizable as a geometric figure. A *terrón* is an orthoedral (maybe, to be simpler, a cube). In cases like this one Form is supposed not to be conceptualised in terms of dimensions as it would be a geometrical calculus, far away from common speaker's knowledge -namely, linguistically irrelevant.
Consequently, we think as more appropriated describing its Form better in a gestalt-style than by means of [JACKENDOFF 1991]'s formalism; this way, we hypothesise that speakers construe a terrón as an object with a cub's shape of a certain size.

In this kind of 'Modelled' portions of masses Quantity must be expressed in absolute (not partitive) magnitudes since it is not pertinent to express it with relation to the whole they are portions of, as long as the whole is indeterminately large -in fact, all of the sugar in the world. For these reasons, in terrón the value of its Size coincides (it is co-indexed) with the value of its Quantity.

---

**Representation of TERRON**

**CATEGORY / STRUCTURE**

```
NP
   /\  \
Q/NP case marker NP
   /\    /\ \
N "de" N

"terrón" x ← type: solids
         99% SUGAR
```

**SEMANTICS**

- quantity: 
- form: SHAPE: cube
- SIZE: ±1 cm³
- agentive: modelling

---

**figure 11.**- lexical entry of TERRON

**[fig 12]** migaja. Migajas haven't got a precise form. As well, they are such small portions that in any way their form is linguistically relevant; consequently we can characterise it by means of the value which we keep for minimal indivisible features: e. Obviously, the same value bears its feature Quantity: a migaja of bread is the most similar thing to no bread at all (even it exists in the language the idiom *ni una migaja*, meaning "nothing").

*Migaja* seems to belong to the class of Detached portions, as it is a small portion detached from the whole (usually bread); but it could also be conceived as Pre-Segmented, if one entertains a bread as something pre-composed of thousands of small *migajas* (the attribution of a basic-level value, e, seems also to suggest *migaja* should be included in the Pre-Segmented class). In this case it would be necessary to discard it to bear the Agentive Role feature, inasmuch Pre-Segmented portions do exist before to be detached from the whole. The existence in the language of the portion noun *pizza*, which also denotes a minimal portion of something but which also straightforwardly denotes the portion has been obtained by detaching it seems to support this alternative view -*migaja* and *pizza* would be equivalent terms but the later would be Detached and the former Pre-Segmented.

**[fig. 13]** extremo. This is the paradigmatic case of Boundary portion, which is equivalent to the English term "end" used by [JACKENDOFF 1991]. It selects solids which could be
conceived as something similar to a line, or at least to something with a strongly relevant longitudinal axis. Its Form is conceptualised as a minimally thick cross cut of the whole.

Representation of MIGAJA

![Diagram of MIGAJA's category/structure and semantics]

**SEMANTICS**
- quantity: $e$
- form: $e$
- agentive: detachment (?)

*Figure 12.* lexical entry of MIGAJA

Representation of EXTREMO

![Diagram of EXTREMO's category/structure and semantics]

**SEMANTICS**
- quantity: $\square$
- form $\square$: DIM 1&2: a cut across the axis of $X$
  - DIM 3: $e$

*Figure 13.* lexical entry of EXTREMO
[fig 14]. *gajo*. This is the paradigmatic case of a Pre-Segmented portion, where the portion did coexist with the whole -thus it does not bear Agentive Role- and whose form is complex and stored in the memory in a non-schematisable way. In the same way, its Quantity is fixed in advance -it does not depend on fuzzy operations of partition but instead, we hypothesise, it is conceived in a basic-level way.

Pre-Segmented portions have a sense of entitativity which is quite stronger than that of other portions; for this reason we think that their Form and Quantity values are conceptually more tied to portions themselves than to the whole they are a portion of. Thus we find sensible representing their Form and Quantity features as referred to the portion itself, not to the whole.

[fig 15]. *jarra*. *Jarra* is a clear example of Container. As a Container portion it selects nouns belonging to [JACKENDOFF 1991]'s category "-B": masses (preferably liquids) and plurals (a glass of water / a bowl of nuts).

Its Quantity refers to the container which the Portion-Term is a methonymical extension of.

*Jarra* shows the interesting issue discussed in Section 5 of the lexicalisation of the Whole, i.e. its specialisation on denoting a portion or a quantity of a specific class of stuff -beer.

For clear cases of lexicalisation such as this one we posit setting two alternative lexical entries: one as an ordinary Container-Portion-Term and another one as a common Noun integrating in its semantics that of the whole it is a portion of -except for Quantity. It is conceivable to obtain this alternative entry by means of a lexical entry expressing a sense extension of the basic entry, as it is discussed in Section 7.2 [F]. But this is an issue we are not going to deal with here.
RELEVANCE FOR NATURAL LANGUAGE PROCESSING.

9. A detailed representation of the lexical entries of partitive terms is of major importance for several aspects of linguistic engineering. The complex conflation of meaning components that these words display is bound be an important source of translation mismatches. For instance, as it can be noticed by looking at (LIST 1), many of them have not a precise word-to-word mapping from Spanish to English, and the same or greater effects are to be expected between Spanish and other even less similar languages. Moreover, the mismatching is brought about for different words by different causes, namely different meaning components. All these factors need to be taken into consideration if a concept-based enterprise of machine translation is to be carried on.

To put some examples, in (25a-d) the English term 'slice' ought to be translated to several different Spanish terms depending on the more fine-grained range of selectional restrictions of Spanish for this area of language. In (25e,f) a single Spanish portion-word need to be translated to a complex group in English, for the Spanish terms conflate in them some adjectival predications which are not denoted by their English counterpart. In (25g,h) the effects of lexicalisation of the whole are shown: while in Spanish una jarra implies by default the content is beer, the symmetric effect is achieved by the English term 'a pint'; in both cases translation of the terms to the other language need of explicit specification of the stuff one is talking about.

(25)  a. a slice of watermelon $\leftrightarrow$ una raja de sandía
      (a slice of watermelon)
b. a slice of ham \(\rightarrow\) una loncha de jamón  
(a slice of ham)

c. a slice of bread \(\rightarrow\) una rebanada de pan  
(a slice of bread)

d. a slice of orange \(\rightarrow\) una rodaja de naranja  
(a slice of orange)

e. un escalope de ternera \(\rightarrow\) a fried slice of beef  
(a fried-slice of beef)

f. córtalo en rodajas \(\rightarrow\) cut it in round slices  
(cut-it in round-slices)

g. beberé una jarra \(\rightarrow\) I'll drink a mug of beer  
(I-FUT-drink a mug)  
\{understood: beer\}

h. a pint \(\rightarrow\) una pinta de cerveza  
(a pint of beer)

An accurate account of the semantics of portion terms is also necessary in order to predict changes in the semantics of nouns thus modified, as for instance:

- in recovering of lexical information from dictionary definitions as reported in [VOSSEN & COPESTAKE 1994] (see discussion above), as long as the semantics of the nouns defined by means of kernels [[PORTION-TERM] of [NOUN]] must be a function of both the semantics of the Portion-relation and the semantics the genus.

- they might cause changes of semantic type involving selectional restrictions of verbs -namely, you can't bend iron but you can bend a sheet of iron; you can drown on water but you can't drown in a glass of water.

- as it was shown in the examples (16) of Section 4, the type shifting of the Agentive Role with respect to that of the whole will play a decisive role in parsing strategies considering the notion of co-composition [PUSTEJOVSKY 1991] which, roughly, posits that the semantic structure of the noun co-acts with that of the verb to get a full interpretation of the phrase.

- as it is pointed in Section 8 (vi), a stochastic representation of selectional restrictions of Portion Terms should be useful for tasks involving decisions on meaning (ambiguous constituent attachment, sense discrimination, ellipsis solving).

**A SPECIAL CASE: PORTIONS OF ABSTRACT CONCEPTS.**

10. About the appliance of portion terms on abstract concepts, such as (26), we will not consider the case here as we assume it to be an episode of the much wider phenomenon of conceptualisation of abstract terms by means of metaphorisation of concrete terms [LAKOFF & JOHNSON 1980]. Namely, it is not a matter of portion terms which may semantically select abstract terms, but instead a more general matter of abstract words conceptualised by means of spatial terms (e.g. (26a) a movie = an oriented line; (26b) mental constructs = graspable solids; (26c) irony = a liquid which can distil from discourses; (26d) arguments = buildings with
foundations, base, solidity, etc.). Consequently, it is a pretty distinct affair and it should be
dealt with elsewhere.

(26) a. el final de una película
    (the end of a movie)

b. un puñado de buenas ideas
    (a handful of good ideas)

c. una gota de ironía
    (a drop of irony)

d. la base de su argumentación
    (the base of his/her argumentation)

CONCLUSIONS

11. Several approaches to the representation of portions in a lexicon have posited linking them
to other nouns by means of relational links such as Portion-Mass, Segment-Whole or so (e.g.
slice-pie, parcel-land, grain-rice). We have shown this is not a fine strategy of representation
inasmuch Portion Nouns are not nouns as the rest but special operators on nouns. Consequently they must be represented in a lexicon as a class separated from that of overall nouns (not embedded in and linked to them) bearing their own sort of linguistic content
-leaving to parsing components of NLP systems the task of construction of syntagmatic
interpretation. In other words, "slice" is not a portion of "pie", what actually is a portion of
"pie" is "a slice of pie", thus the very different kinds of semantic content of "slice" and "pie" is
what should actually be combined to get the semantics of the expression.

Portion Terms are complex words which share properties of referential expressions
(nouns), operators (quantifiers), predicates (adjectives) and specifiers (determiners). Their
complexity is revealed by the following considerations:

(i) they don't behave as nouns in that they are not heads of NPs but heads of operators adjunct
to head-NPs; nevertheless, they are categorically Nouns and they head their own operator-NPs;

(ii) they compulsory govern nominal structures in genitive, case-marked in Spanish by the
preposition de; but in some cases they appear without complementation thus lexicalising the
semantic content of the complement;

(iii) they are quantifiers as they specify a quantity or measure of the whole;

(iv) they are specifiers in the sense that the portion is attributed a sense of entitativity -even
applying on masses portions turn to be individuated objects;

(v) they are supposed to combine with mass terms, but they also combine with individuated
objects; specially some of them -portions denoting boundaries such as "end" or "tip"- can
 overtly select individuated objects;

(vi) they bear properties of referential nominals, as the Agentive Role, thus the claim of full
transitivity of properties between wholes and their portions has to be rejected;

(vii) they act as predicators (more precisely as adjectives) in that they predicate properties of the
nouns they combine with, as shape, size or other;
(viii) they also behave as other predicates (verbs, adjectives) on constraining the kind of nouns to combine with —namely, they bear selectional restrictions.

A taxonomy of Portion Nouns, grounded on cognitive distinctions but intended to map on a featured hierarchy ordered by subsumption has been outlined. It results on the following main classes: Container Portions, Detached Portions, Modelled Portions, Pre-Segmented Portions and Boundary Portions: Container Portions bear no Form of their own but the other do; Detached and Modelled Portions did not exist before to be separated of the whole by some process (detachment or mass-modelling), Pre-Segmented and Boundary Portions did; Pre-Segmented Portions are units of a naturally pre-existing internal structure of the whole; Boundary Portions are conceptualised as minimal elaborations of ideal limits of the whole.

Cognitive Grammar notions of Image-Schematic and Basic-Level concepts and Jackendoff's way of conceptualisation of parts and boundaries seems to us to be appropriate to represent the speaker's knowledge of Portions Terms. For instance, *rebanada, lámina, oblea* or *rodaja* are hypothesised as conceptualised as planes; *tira* or *segmento* as lines; *punta* or *extremo* as points.

In terms of features, they mainly diverge on different sorts of content for Form, Quantity, Agentive Role or Selectional Restrictions: Container Portions express the Quantity naturally expressed by the container they are metonymies of and exclusively select masses and plurals; Detached and Modelled Portions bear Agentive Role (detachment and modelling, respectively), and their Quantity and Form are conceptualised image-schematically; Pre-Segmented Portions select solids with internal structure and their Form and Measure are globally (basic-level) conceptualised; Boundary Portions select individuated objects and their Quantity and one of their Form-dimensions are conceptualised as minimal.

**FURTHER WORK**

**12.** Indeed this is a general approach to the issue and lots of work remains to be done, specially:

- The fact that portion terms sometimes don't appear complementing a noun but stand alone having lexicalised the Whole in them (these are the cases of *quinto* and *jarra*) is repeatedly mentioned along the paper. It might be thought it is plausible predicting somehow this behaviour by means of lexical rules. The case is similar in the surface performance to that of the ellipsis of the whole *filete* usually meaning *filete de ternera*, see footnote 1; which in turn is related to assigning weights calculated from co-occurrence in corpora to nouns selected by portion terms; the noun with the higher weight may be the best candidate to solve the ellipsis. But perhaps this is not the case as long as the default selection may be so frequently elided that it may not be the higher-weighted co-occurrence. All these problems deserve further accurate attention.

- The appliance of Portion Terms—which seem naturally intended to deal with concrete things-on abstract objects is a field to be explored. It seems that a broad treatment of metaphorisation is to be involved.

- Taxonomies and representations in the paper are outlines—a though work on precise formalisation to be used in NLP and knowledge representation is to be carried on. Types for the classes and for their values should be defined in terms of feature structures.

- In general, Portion Terms mentioned need deeper account; other which have not been specifically treated need to be dealt with; and eventually new terms or classes of terms which
look plausible to fit in the framework have to be found and integrated in it.

- Finally, some verbs are directly related to portions, e.g. *trocear* [to cut up into pieces] and *rebanar* [to cut up into *rebanadas*] - they seem to be either derivates of portion terms or else portion terms derive from them. These verbs show how different types of processes bring about different types of portions. Relations between such verbs and nouns of portions, - involving the notion of Agentive Role- also should be considered.
REFERENCES:


JACKENDOFF R. 1989. What is a Concept, that a Person May Grasp It? Mind & Language, 4, 68-102


38


LSI-95-1-R "Octree simplification of polyhedral solids", Dolors Ayala and Pere Brunet.


LSI-95-3-R "The complexity of searching implicit graphs", José L. Balcázar.

LSI-95-4-R "Design quality metrics for object-oriented software development", Alonso Peralta, Joan Serras, and Olga Slavkova.

LSI-95-5-R "Extension orderings", Albert Rubio.

LSI-95-6-R "Triangles, ruler, and compass", R. Juan-Arinyo.

LSI-95-7-R "The modifiability factor in the LESD project: definition and practical results", Nuria Castell and Olga Slavkova.

LSI-95-8-R "Learnability of Kolmogorov-easy circuit expressions via queries", José L. Balcázar, Harry Buhrman, and Montserrat Hermo.

LSI-95-9-R "A case study on prototyping with specifications and multiple implementations", Xavier Franch.

LSI-95-10-R "Evidence of a noise induced transition in fluid neural networks", Jordi Delgado and Ricard V. Solé.


LSI-95-12-R "Computer (PC) assisted drawing of diagrams for forecasting soaring weather", Lluís Pérez Vidal.

LSI-95-13-R "Animats adaptation to complex environments as learning guided by evolution", Mario Martin, Màrius Garcia, and Ulises Cortés.

LSI-95-14-R "Learning to solve complex tasks by reinforcement: A new algorithm", Mario Martin and Ulises Cortés.


LSI-95-16-R "MDCO to B-Rep conversion algorithm", Dolors Ayala, Carlos Andújar, and Pere Brunet.
LSI-95-17-R “Augmented regular expressions: A formalism to describe, recognize, and learn a class of context-sensitive languages”, René Alquézar and Alberto Sanfeliu.


LSI-95-20-R “From degenerate patches to triangular and trimmed patches”, Marc Vigo, Núria Pla, and Pere Brunet.


LSI-95-23-R “Approximating the permanent is in RNC”, J. Díaz, M. Serna, and P. Spirakis.

LSI-95-24-R “Constraint satisfaction as global optimization”, Pedro Meseguer and Javier Larrosa.


LSI-95-26-R “External schemas in object oriented databases” (written in Spanish), José Samos.


LSI-95-28-R “Dynamic belief modeling”, Antonio Moreno and Ton Sales.


LSI-95-30-R “Difference lists and difference bags for logic programming of categorial deduction”, F. Xavier Lloré and Glyn Morrill.

LSI-95-31-R “An experiment on automatic semantic tagging of dictionary senses”, German Rigau.

LSI-95-32-R “Slices of meaning: a study on nouns of portions”, Salvador Climent and Maria Antònia Martí.

Hardcopies of reports can be ordered from:

Nuria Sánchez
Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Pau Gargallo, 5
08028 Barcelona, Spain
secrelsi@lsi.upc.es

See also the Department WWW pages, http://wwwlsi.upc.es/www/