Pirates, Who are They?
A Cognitive-Linguistic Analysis of Italian Media Discourse

ANNARITA GUIDI
Roma Tre University

The aim of this article is to propose an outline of piracy as it is represented in the Italian Web press. Departing from the idea that media relationships are conceived as bounded in contractual and commercial “mental” spaces, I analyze the correlation between the semantic structures of pirateria (“piracy”), pirata (“pirate”), and the use of such words in two Web corpora. Results show that the linguistic correlates of “piracy” conceptualization can be explained in terms of image schemas, which recur in metaphorical patterns and syntactic-semantic frames associated with “piracy” and “pirate.”

Introduction

At the present time, seen from the Italian Web press, the phenomenon of piracy seems to hang in a delicate balance. According to a research of NpD Group, the actions directed to important file-sharing websites (e.g., LimeWire) have brought about a sensible decline (from 16% to 9%) in the percentage of people whose habitual access to cultural content takes the shape of peer-to-peer content sharing. At the same time, the Web press reports the constant growth of other portals (e.g., Frostwire, BitTorrent) which

---

1 I wish to thank Maria Catricalà, because without her insightful reflections and sustain, this article would not have been written. I am very grateful to the people who supported me with their expertise and careful suggestions: Raffaele Altieri, Gianluca Altieri, Federico Chiacchiari, Rony Medaglia, Sergio Sozzo, and Varo Venturi. Also, thank you to the anonymous IJoC reviewers for their valuable input.
serve today the 33% of people who “illegally” download music files, while P2P traffic volume is “higher than all others together” in many regions of the world (Cardoso, Espanha, Jacobetty, & Lima, 2009).

Perhaps the first impression that one gets while having a look at media discourse on piracy is the kind of confusion which characterizes the transition toward horizontal communication networks. In this article, I propose a cognitive-linguistic analysis of the “intellectual categories” (Castells, 2000, pp. xvi–xxx) which designate a cluster of phenomena, including piracy, peer-to-peer file sharing, and counterfeiting. Blurring the boundaries between mass media and other forms of communication, making “literally anything” susceptible to being “digitized,” such a transition is supposed to have an impact on the patterns we use to represent communicative events.

From a sociological perspective, such patterns have been defined as “mass self-communication” and, considering the “malleability” of Internet as a technology (Castells, 2001, p. 5), they appear to be closer to a Piracy Manifesto than the (Italian) copyright laws. At the same time, the representation of piracy in the Italian Web press still appears to be an ideologically-based conceptualization of events. That is, specific ideologies superimpose (also through the use of metaphors in argumentative texts; see Partington, 2006) an evaluation on the discourse’s referents, acting as interfaces between the representations underlying media discourse and the interests of specific social groups.

From the Sociocognitive Approach to Space Grammar: Context Models, Imagery, and Prominence

In order to outline the conceptualization which underlies the Italian Web media discourse on piracy, a non-traditional definition of context is assumed. According to van Dijk (2008), context is not only determined by independent social variables (e.g., age), but it is also a mental model—a subjective interpretation of communicative situations. For our purposes, the relevant corollaries of this theory are the following: (a) “it is not the social situation that influences . . . discourse, but the way the participants define such a situation” (ibid., pp. x–xi); (b) texts and talks not only are constituents of their contexts, but also are constitutive of them; (c) context models consist of schemas of shared, culturally based, conventional categories and remain largely implicit and presupposed; and (d) context models control how participants produce and understand discourses and events.

How can a context model influence the lexicon of a language? Since such models define “how we see” a situation, they can be compared to the “projected world” (Jackendoff, 1983, pp. 17–32). The author’s hypothesis is that “there is a single level of mental representation, conceptual structure, at which linguistic, sensory, and motor information are compatible.” As a consequence, “semantic structures could be . . . a subset of conceptual structures—just those conceptual structures that happen to be verbally expressible.” The point is, “one cannot perceive the real world as it is . . . [P]otentially vast areas of our experience are due to the mind’s contribution, even though the experience is of things ‘out in the real world.’ . . . The projected world embraces not only direct perceptual experience . . . but also a wide

---

2 Within the cognitive theory of text processing, such a definition of context provides the missing link between mental models of events talked about and the way in which discourse is formulated.
variety of abstractions and theoretical constructs.” The “reference” of linguistic expressions, thus, is not the real world, but the “projected world.”

Since the context model is a social, dynamic construct of the participants about the “for-them-relevant properties of a situation” (van Dijk, 2008, p. 56), it is also compatible with Langacker’s (1987, 1990, 2009) prominence-based model. Cognitive (originally space) grammar considers each linguistic datum as a symbol, e.g., a holistic, non-compositional entity made up of correspondences between semantic components, phonetic components, and their sub-structures. Semantic structures are the conceptual structures evoked by linguistic expressions. There are many ways of shaping a meaning: In terms of cognitive grammar, this means that a verbal expression (such as piracy) imposes a particular “image” on the content it evokes. The prominence-based model sees the “imposition of a profile on a base” (Langacker, 1990, pp. 5–15) as the first dimension of imagery; meaning is given by the selection of a particular substructure within the base.

In terms of communicative purposes, using a particular expression corresponds to selecting a particular image, which structures a situation. An example will clarify the method used here to analyze piracy-related expressions. The sentences (a) Bill sent a walrus to Joyce and (b) Bill sent Joyce a walrus differ in meaning, because they employ different images to structure the same event. In (a), the morpheme to designates the “path” followed by the walrus, rendering this aspect of conceptualization more prominent. In (b), the juxtaposition of the nominals (Joyce; a walrus) symbolizes a “possessive relationship”—the prominent configuration is that in which the walrus has completed its “trajectory.” Such difference in imagery has an impact on the “felicity” of using to instead of the double-object construction for certain types of situations. For example, since a continent is hardly susceptible to being construed as a possessor, a sentence like I sent Antarctica a walrus is felt to be less acceptable than I sent a walrus to Antarctica.

---

3 A base is the domain of a predication. A domain is “any sort of conceptualisation: a perceptual experience, a concept, an elaborate knowledge system, etc.” Examples of basic domains are the experience of time and the spatial configurations. The profile of a base is a “substructure elevated to a special level of prominence within the base” (Langacker, 1990, pp. 5–15).
4 As in Langacker’s model of analysis, heavy lines in Figures indicate a certain degree of relative prominence.
Along with imagery, a further methodological tool relevant for our purposes is the concept of schema. We can distinguish between schematic units and image schemas. The first ones represent the content shared by conventionalized units and image schemas. The first ones represent the content shared by conventionalized units; e.g., thing is a schema instantiated by lexical units such as pencil or fork, which are individual members of the schema. Categorizing processes reflect the judgment that a specific member instantiates the schema. An image schema is "a recurring dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience. . . . 'Experience' is to be understood in a very rich way, including basic perceptual, motor-program, emotional, historical, social and linguistic dimensions" (Johnson, 1987, pp. xiv–xvi; see also Lakoff & Johnson, 1980; Lakoff, 1987). The basic inventory of image schemas includes the **container/containment** schema and the
The Lexicon of Piracy: Mappings, Etymologies, Image Schemas

As we have seen, context models represent the interface between mental models of events and the discourse on such events. This study departs from the hypothesis that the mental model of events related to piracy is regularly structured by two image schemas—a regularity reflected in language use.

As pre-conceptual and spatial structures, image schemas represent the basis of metaphorical mappings (Lakoff & Johnson, 1980). Etymological data show that the concept “illegal copying” founds the metaphorical meaning of piracy. The crucial point of the mapping illegal copying is piracy is the relationship between such concept and that of “appropriating something,” included both in the literal and figurative meaning of piracy. Establishing a connection between “copying” and “appropriating” corresponds to making prominent specific aspects of the mapping, and to clouding other aspects. The latter ones become the “used part” of the piracy metaphor in some discourse domains, such as cinema and politics (see below). The prominent aspects form a mental model and correspond to two image schemas. This can be seen through the aggregation of corpora data—metaphors, metaphorical patterns, multi-word expressions, noun phrases (head-modifier) and predicate-arguments structures—around the image schemas source-path-goal and container/containment.

Image schema analysis shows that (a) “illegal digital copying” inherits specific properties of “piracy” and comes to be represented as an attack and appropriation, and (b) that the context model of “counterfeiting” is different from that of “illegal digital copying.” Data aggregation also shows the coherence relationship (Lakoff & Johnson, 1980, pp. 64–65) among metaphors such as illegal copying is piracy, illegal copying is fishing, and illegal copying is killing, which share the implication illegal copying is appropriating something. Image schemas embedded in the linguistic representation of “piracy” let some problems related to the difference between the “piracy” context model and the Web users’ actual practices.
also emerge. Such problems involve the legal aspects of piracy and the concepts of “intention,” “awareness,” and “theft.”

For the purposes of analysis, the relevant categories of the context model are: (a) intention (“the construction of a mental model of an ongoing or future fragment of conduct”; van Dijk, 2008, pp. 81–88), and (b) purpose (“mental models of actions and their wanted consequences”). In terms of shared knowledge, the sociocognitive strategies assumed by journalists can be formulated as follows: “Assume that readers have the same sociocultural knowledge as I have.” From a semantic-pragmatic perspective, the shared knowledge of intentions and purposes represents the basis of the readers’ inferences concerning the non-explicit statements present in a text.

Let us first consider data of lexical variation. Pirateria and pirata appear to be present in all the “situation types” (van Dijk, 2008): formal (including public and institutional situations), standard (e.g., journalism), and informal (e.g., Web forums). This fact emerges from corpora analysis, as well as from the recent hearing of the Italian Antitrust Authority concerning “counterfeiting and piracy” (Catricalà, 2011). Here contraffazione (“counterfeiting”) refers to the selling of goods which constitute illegal reproductions of trademarked or patented products; the intellectual property gives the patentee the exclusive right of commercial exploitation. Concerning pirateria and diritto d’autore (“copyright”), the hearing highlights the fact that online editorial content is used and re-edited by “third subjects” (e.g., search engines), while the editors of such content cannot gain from the “further value” of their own editorial activity. With a farsighted remark, the Antitrust President notes that the Italian copyright laws are not able to handle all the technological and economical dimensions of the Web; “models of virtuous cooperation” among copyrights’ holders and providers of innovative Web services (e.g., aggregators) are thus indicated as hopeful solutions. Looking at the relationship between “consumers” and “goods,” the Authority states that the majority of “counterfeiting and/or piracy” cases can be ascribed to two typical situations: (a) the consumer is aware of (and satisfied with) the fact that the product is counterfeit (merci contraffatte), or (b) the consumer does not even perceive the violation of copyright as an unlawful practice, and the good he/she gets is exactly “the same” of the protected one. That is, from a legal point of view, counterfeiting appears to be different from piracy because of: (a) the consumer’s position (awareness vs. non-awareness), and (b) the product’s nature (different from vs. identical to the “original” one).

We will return to the second point below. Concerning the consumer’s awareness, it is interesting to look at etymological data and the position of pirateria and pirata in the ontology Italwordnet. Since corpora data show an overlap of the piracy, information technology, and web user domains, it is worth to note, first, which are the semantic-pragmatic entailments of the analysed words, and second, whether such entailments do or do not show a correspondence with language use.

---

8 On the historical development of this concept, see Johns (2009).
Pirate: The Schema

The figurative sense of piracy refers to “the unauthorized reproduction or use of a copyrighted book, recording, television program, patented invention, trademarked product” (Collins, 1991, 2009). It also refers to “the unauthorized copying, distribution, or use of another’s production (as a film) especially in infringement of a copyright; the unauthorized use, interception, or receipt of encoded communications (as satellite cable programming) especially to avoid paying fees for use” (Merriam-Webster, 1996). The figurative sense of pirate as “one who reproduces,” “uses,” and also “appropriates” is recorded in 1668 (“Some dishonest booksellers, called land-pirates, who make it their practise to steal impressions of other men copies”; J. Hancock, cited in Onions, 1966). This is also the semantic domain of the verb pirate (Italian piratare, “to rob by the sea,” first recorded in 1598); its figurative meaning (recorded in 1706) is “to appropriate or reproduce the work of invention of another without authority for one’s profit.”

Such data show that, as a lexical entry, piracy is not clearly separated from counterfeiting; that is, piracy also refers to trademarked products and is associated with a profit. A first problem is thus represented by the fact that pirateria and contraffazione are actually conceptualized as different entities, at least in part. Such separation is based on the categories of intention and purpose. A second problem is represented by the fact that an overlap of the meanings of pirata and hacker emerges in Italian, while English seems to maintain an opposition between the two nouns.

Nonetheless, the etymology of pirate and the “semantic shifts” (Simone, 2007) of hacker display a shared semantic area. Piracy (Anglo-Latin pirātia < late Greek peirāteia) originates from pirate (Latin pirāta < Greek peirātēs [“one who attacks”] < peirā [“to attempt, to attack”]; Collins, 1991, 2009), the etymology of which refers to “to attack, make a hostile attempt on, try” (Harper, 2010). As a verb, hack shows the interrelated literal meanings “to cut, to slice,” “to break up the surface (of the ground),” “to damage, to injure.” The figurative sense of “a try, an attempt” (with a shift verb > noun) is recorded in 1898. The conceptual overlap of pirate and hacker is thus related to the meaning of attempt (“attack, aggression”).

As a noun, hack also refers to “a person hired to do routine work” with the special sense of “one who writes anything for hire,” first recorded in 1826. In 1980s, hacker becomes associated with the extended meaning, “one who gains unauthorized access to computer records.” A metonymic referential shift (hired worker > enthusiast computer programmer > person who illegally penetrates digital systems) has occurred, since this meaning is said to be from the earlier slang sense of “one who works like a hack at writing and experimenting with software, one who enjoys computer programming for its own sake.” The verb hack (“illegally enter a computer system”) is first recorded in the same period.

In ItalWordNet (Roventini, Alounge, Calzolari, Maglini, & Bertagna, 2000), pirateria shows the hypernyms atto (“act”), azione (“action”), and moto (“motion”). Every meaning structure associated with pirateria would thus include the dimension of intentionality. In terms of frame-frame relationships (Fillmore,

---

9 Castells (2001, pp. 41–50) argues that hacker is an ambiguous word and hackers “are not what the media say they are,” since they are identified with crackers.
Johnson, & Petruck, 2003), piracy inherits properties from committing crime and intentionally act.\(^{10}\) This is obvious from a semantic point of view (though not from a legal one); pirata shows indeed fuorilegge, criminale ("out of law") as its hypernyms.

On the other hand, contraffazione shows the hypernyms atto ("act"), azione ("action"), copia ("copy"), and riproduzione ("reproduction"); but the hypernym of contraffattore ("counterfeiter") is just individuo ("individual"). Even more important, the near-synonyms\(^{11}\) of contraffazione are alterare and falsare ("to modify, to falsify"), while that of pirateria is prendere ("to take"). If we connect these data to the verbs for which pirateria and pirata play the semantic role of agent (see below), the picture becomes definitely clear.

We can now associate two image schemas with the conceptualization of piracy and pirate: path and container/containment. They profile two different components of the meaning structure. The path schema conceptualizes the (abstract) motion of an agent (the trajector) toward a patient (the landmark). The profiled sub-structure of such motion is the impact of the trajector onto the landmark (source-path-goal), which corresponds to "attempt"—the semantic root of pirate and the figurative meaning of hack. This profiled sub-structure corresponds to the concept of intentionally damaging something. The container/containment schema profiles a specific consequence (purpose) of the path schema; that is, the appropriation (perpetrated by the trajector) of something belonging to the landmark. The path schema profiles the final state (goal) of a process; the container/containment schema profiles the result of the process.

\(^{10}\) See also the verb to pirate as an instantiation of robbing and stealing; the noun pirate as an instantiation of kinds of theft; and the noun piracy as an instantiation of acts of stealing (McArthur, 1981).

\(^{11}\) For an overview of synonymy and a review of the related linguistic tests, see Catricalà (2009).
As we have said, the difference between a hacker and a pirate is blurred in the Italian language use; the border between the meanings of contraffazione and pirateria is not so clearly marked (this can also be inferred from the fact that they are often grouped together at an institutional level).\textsuperscript{12} Who does actually emerge as a pirate in language use?

\textsuperscript{12} For the European Observatory on Counterfeiting and Piracy, see http://ec.europa.eu/internal_market/iprenforcement/observatory/index_en.htm
The theories sketched above find in corpus-based approaches a field of empirical verification through authentic language data. Such an approach allows the study of conceptual metaphors from a linguistic perspective, and it can fruitfully complement the qualitative analysis of texts aimed to determine the ideological, social, communicative, and cultural functions of metaphor (Stefanowitsch, 2006a). Such aspects of metaphors emerge from the ideological selection of expressions and the cognitive models on which discourse is based (Koller, 2006, pp. 237–239). If cognition is “distributed” among the members of a community, mental models can be imported to (and multiplied in) social structures, making even pre-conceptual schemas part of the sociocultural context (Gibbs, 1999). This idea can be verified by analyzing the language use of a large number of speakers, also considering the fact that (Web) journalists are likely to show a high degree of readership orientation, so as to echo and reinforce specific conceptual models.

In order to verify whether the path and the container schemas, emerging from the original semantic domains of piracy, do or do not have correlates in current language use, the analysis takes into account data coming from two Italian Web corpora. The first one is itWac (Baroni & Kilgarriff, 2006; www.sketchengine.co.uk, totalling 1,575,489,232 words). The second one, WoP, is a small, non-representative corpus formed by 40 articles on piracy and counterfeiting that appeared on the Web in March, 2011, when the media discourse on piracy culminated with the report of what has been regarded as a historical landmark: the order of the Court of Rome involving the media company Yahoo! (see Appendix). As a small data set, WoP serves as a qualitative basis to study how specific verbal expressions can reflect the ideology of participants in the discourse on piracy. These results are compared with language patterns retrieved from itWac. A further difference concerns the fact that WoP contains only articles, while itWac includes any kind of source, with an important portion coming from blogs and forums (it can, thus, be considered as a “computer-mediated communication” corpus. On this issue, see Orletti, 2004).

Some scholars (e.g., Ponterotto, 2005, p. 157) highlight the necessity of linguistic inquiries based on an integration of cognitive theories and discourse analysis and aimed to study the role of conceptual metaphors in communicative processes, e.g., the role which metaphoric conceptualization plays with respect to the comprehension and production of discourses seen as “context-bound events.” Ruiz, de Mendoza Ibanez, and Velasco (2005, p. 170) see as a “paradox” the rarity of empirical studies (based on authentic, ecological data) within cognitive linguistics. Kövecses (2011, pp. 24–33) highlights the fact that the traditional approach to metaphor does not pay attention to “which actual metaphorical expressions are used . . . by real speakers in natural discourse.” The author proposes a corpus-based methodology of identification of metaphorical expressions, which includes the study of their syntactic, semantic, and pragmatic behavior in “concrete contexts of use.” One important goal is “to see to what extent the metaphors contribute to the conceptualisation of abstract concepts, as well as their cognitive representation, and what content they contribute.”

For an overview of representativeness of corpora, see Lenci, Montemagni, and Pirrelli (2005).
The analysis takes into account the following categories:

- **MWE_p**: MultiWord Expressions whose components correspond to *Pirateria*, *Pirata*, *Pirati* (e.g., *pirateria digitale*, "digital piracy");

- **MWE_s**: MultiWord Expressions whose components are used in the text as synonyms (or periphrases) of the previous ones (e.g., *scambio illegale*, "illegal exchange" for *pirateria*);

- **M_p**: Metaphors which represent the concepts “piracy,” “pirate” instantiating an image schema (e.g., *esploratori*, “explorers”), or a conceptual mapping (e.g., *nemici dei distributori*, "enemies of the distributors");

- **M_r**: Metaphors and Metaphorical patterns which represent concepts related to piracy (e.g., *argini di difesa*, literally “dike for defense,” “barricade”).

A first result is the overlap of the conceptual domains of *pirateria* and *informatica* ("information technology"). The MWE_p and MWE_s retrieved from the corpus show that *piracy* is often identified with *computer piracy*, as can be seen in the following examples.

MWE_p: *pirateria digitale* ("digital piracy"), *pirateria in rete* ("Web piracy"), *pirateria on line* ("online piracy"), *pirateria software* ("software piracy"), *pirateria su pc* ("computer piracy"), *build pirata* ("pirate build"), *portali pirata* ("pirate Web portals"), *programmi pirata* ("pirate computer programs"), *siti pirata* ("pirate Web sites"), *software pirata* ("pirate software")

MWE_s: *coloro che scaricano film dalla rete* ("people who download movies"), *cracking*, *file hosting*, *frequentatori delle reti Torrent* ("people used to access Torrent networks"), *hacking*, *link illegali* ("illegal links"), *peer to peer*, *popolo del file-sharing* ("file-sharing people"), *visione in streaming* ("streaming").

That means, the prominent sub-structures of the concepts “pirate,” “piracy” ("to impact on something and damage it"): **path**; “to take (away)”** containment** are put in relationship much more with actions perpetrated on the Web and in multimedia content, than with actions perpetrated elsewhere, with other kinds of products (as happens in counterfeiting). The difference between the concepts "file sharing" and "piracy" tends to be clouded; *pirata* and *illegale* ("illega") are used as synonyms. This first

---

15 Stefanowitsch (2006b, p. 66) proposes the following definition of metaphorical pattern: “a multi-word expression from a given source domain into which one or more specific lexical items from a given target domain have been inserted.” Such a subclass of metaphors, thus, “contain[s] both source and target domain lexemes.” In *argini di difesa*, the source domain is **containment** (whether natural or artificial, e.g., *argini di fiume*, literally “dikes of a river”; *argini di una pianura*, literally “dikes of a plan”; or *argine di cemento*, "cement dike"). The target domain is that of **war**.

result is validated through itWac. Non-compositionality of MWE also reflects the fuzzy boundaries among legal and non-legal domains of action. For example, the MWE_s scambio illegale di film ("illegal film sharing") reflects the difficulty of determining which is the meaning of illegal in such a construction. File sharers violate the copyright at the same time that they exert a fundamental right. The non-compositionality of such expressions has to be intended from this point of view, which is, in some way, validated by the following excerpt (article 12):

La decisione della Paramount mette in evidenza come la rete BitTorrent possa essere utilizzata senza infrangere in alcun modo le leggi in materia di copyright, distribuendo contenuti liberamente accessibili con tutti i vantaggi che l'utilizzo di tale protocollo possiede.

[The Paramount decision highlights the fact that the BitTorrent network can be used with no infringement of the copyright laws, through the distribution of free and open contents, with all the related advantages.]

A second result concerns the overlap of pirates and Internet users, which can be seen in MWE_s such as the most part of film consumers, the most part of DVD consumers, filmgoers, Internet users, and also peer to peer. Such overlap is obviously related to the previous one and provides linguistic evidence of the spread of "illegal" practices (such as the downloading of multimedia content) that leads to a representation of the Internet user as a pirate. Such spread is, on one hand, a problem that must be handled at the legal level; on the other hand, it does not attenuate the negative valence associated with piracy but not to other practices (even those that are very similar in terms of a damage perpetrated over the property rights' owners).

A look at the metaphors and metaphorical patterns used in the Web press shows another two results. First, expressions which instantiate the conceptual metaphor legal action is war are used with respect to computer piracy only—they never "evaluate" (see above) the actions which institutions make with respect to other phenomena (such as counterfeiting). This seems to be due to the resonance between the conceptualization of piracy and that of war: if pirates negatively impact things (e.g., intellectual property, copyright, rights of commercial exploitation), damaging them and also taking away things to which they have no rights, it is clear that they will be considered (at least in a specific ideological view) as

---

16 Such a complex question cannot be exhaustively discussed here, but it is worthwhile to note that the anonymous practice of sharing files can be ascribed to a set of fundamental rights (e.g., freedom of expression and access to culture). At the same time, it is recognized as an illegal practice, even if individuals tend not to perceive the unlawfulness of their own behaviors.

17 Web users is considered a more apt description than consumers, since the "Web democratization" would not be linked to the vast number of "common-people" with access to the Internet, but rather to the greater number of those who freely elaborate, download, and upload written and audio-visual information, opinions, and entertainment content to and from the Web (Martín-Barbero, 2009, pp. 808–809). Since P2P networks’ users are both enablers of, and participants in, a content and technological resource sharing process, they can also be considered “distributors” (Cardoso, Espanha, Jacobetty, & Lima, 2009).
enemies. Other metaphors seem to be coherent with the conceptualization model proposed above: Pirates are represented as fishers who cast hooks in the sea of the Web, an expression which instantiates the image schema source-path-goal, profiling the goal (e.g., the impact on the surface of what will be entered). Other metaphors, such as capture of codes and trap e-mail instantiate the schema container/containment, profiling the second term (e.g., the appropriation of something).

A further difference between “illegal digital copying” and “counterfeiting” emerges through the selections that speakers perform on the ground of their context model. The places in which pirates are “at work” are called centrali, centri (“centers”) and snodi (“junctions”), while the counterfeiting places are called case (“houses”). On one hand, this can be due to the fact that counterfeitters are conceptualized as actual producers of things, while the common representation of digital piracy may entail the fact that things arrive to (e.g., downloading) and depart from (e.g., uploading) pirates’ places. On the other hand, in terms of conceptualization, this difference can be ascribed to a dynamic force (e.g., a movement from/to piracy centers) embedded in the meaning structure of piracy, which is coherent with the path schema (the domains of which are space and time) and absent in the meaning structure of counterfeiting (casa is conceptualized only within the domain of space).

A Large Corpus: Piracy and the Italian Web

As shown by van Dijk (2008), usage-based models can enlighten, at least in part, not only the contexts in which words are used, but also the socially shared conceptualizations associated with them. In itWaC, the frequency (number of tokens) of pirateria is 7.033 (0.00035%). Despite this low level of frequency, the corpus provides some interesting syntactic-semantic clues which seem to support the idea of a specific, dominant, widespread conceptualization of piracy.

Let us consider first the overlap of the domains of piracy and information technology (see above). Looking at noun phrases, the list of adjectives associated with the noun pirateria starts with informatica; the salience value of such construction is the highest one18 (7.32; 725 tokens). Informatica is followed by audiovisiva (“audiovisual”), which still specifies the commercial/industrial domain of pirateria, and by dilagante (“spreading”; salience 6.84). Other adjectives defining the commercial/industrial domain are musicale (“musical”; 5.92), on line (5.27), and multimediale (“multimedia”; 5.24). Another interesting datum is represented by the and-or pattern (instantiated by expressions such as X and Y, X or Y), in which pirateria is mainly associated with the words copia (“copy”), copyleft, contraffazione (“counterfeiting”), cyberdelinquenza (“cyber-criminality”), downloading, falsificazione (“falsification”), file-sharing, hackeraggio (“hacking”), hacking, and imitazione (“imitation”).

The verbs for which pirateria plays the syntactic role of subject (frequency: 1.040) entail a (physical or not) damage to the object: danneggiare (“to damage”; incidentally, in ItalWordNet this verb shows pirata as an instantiation of the category involved agent), distruggere (“to destroy”), falcari (“to

---

18 "Salience" has to be intended as a measure of the degree of lexical association between two words (the noun and the adjective in this case). Salience is based on several factors (e.g., frequency of the words, and the probability that they occur together).
mow down”), flagellare (“to flagellate”), intaccare (“to nick”), invadere (“to invade”), rovinare (“to ruin”), and uccidere (“to kill”). It is worthwhile to note that such verbs do not automatically entail the semantic role of agent. This fact is related to the shift pirata > pirateria, which encodes a lower grade of agentivity correlated to a lower grade of “concreteness” (Sadoski & Paivio, 2001) and “nouniness” (Ross, 1973) of the noun itself. Verbs associated with piracy instantiate a path schema profiling the goal; the contact between the trajector and the landmark is symbolized as a damage. For verbs such as uccidere, the schema interacts with the container/containment one19 (e.g., to kill is conceptualized as “to deprive someone of his/her life,” or “to take someone’s life”). Figure 5 synthesizes these results and includes the hypernyms shown by ItalWordNet.

![Diagram](image)

**Figure 5. Verbs associated with “pirateria”: Path and container/containment image schemas.**

---

19 Many studies provide linguistic data on the combination of these schemas. See, for example, Hampe and Grady (2005).
The verbs for which pirateria plays the syntactic role of object (frequency: 740) validate the war conceptual metaphor as the most used one; the dominant representation of piracy is that of something to be fought. 365 tokens are distributed among verbs such as abbattere (“to tear down”), combattere (“to fight”), contrastare (“to contrast”), debellare (“to defeat”), eliminare (“to eliminate”), reprimere (“to repress”), sconfiggere (“to overcome”), or stroncare (“to strike down”).

Let us now have a look at the noun phrases containing pirata (frequency: 14,513; 0.00072%). The highest level of salience corresponds again to the adjective informatico (7.31; 732 tokens). At lower levels of salience, we find the modifiers cyber (4.67), cibernetico (“cybernetic”; 3.21), musicale (“musical”; 3.14), telematico (“telematic”; 2.96), editoriale (“editorial”; 1.06), on-line (0.05), and elettronico (“electronic”; 0.02). In and-or patterns, pirata is (not surprisingly) associated with hacker (6.6), which represents the first relevant datum in the list; the only others are cybersquatter (4.96), truffatore (“cheater”; 4.57), file-sharer (3.55), fuorilegge (“out of law”; 2.56), and fruitore (“user”; 1.72).

Looking at verbs for which pirata plays the syntactic role of subject (frequency: 1.966), we find a series of expressions which entail a motion: assalire (“to assail”), assaltare (“to assault”), attaccare (“to attack”), colpire (“to hit”), entrare (“to enter”), infestare (“to infest”), penetrare (“to penetrate”), and violare (“to violate”). All these cases instantiate a path schema in which the contact (goal) between the trajector and the landmark is profiled; some of them (see Figure 6) also instantiate the goal through the semantic sub-structure corresponding to “damage.” Another series of verbs entails the actions of “taking possession” or “taking (away): approfittare (“to take advantage of”), derubare (“to rob”), duplicare (“to duplicate”), impadronirsi (“to take possession”), impossessarsi (“to take possession of”), saccheggiare (“to plunder”), sfruttare (“to exploit”; in ItalWordNet, pirata is an involved agent for this verb), and truffare (“to cheat”). These verbs instantiate the interaction between the two schemas (path and container).

---

20 ItalwordNet and Framenet do not encode (e.g., as a purpose or a result) the meaning “taking possession” in the ontological maps of the verb duplicate. This is the reason why I have not included it in Figure 6.
Figure 6. Verbs associated with "pirate": Path and container/containment schemas.

The verbs for which pirata plays the syntactic role of object (frequency: 780) show a picture very similar to that of pirateria: 102 tokens distributed among arginare ("to contain"), catturare ("to capture"),
Combattere ("to fight"), condannare ("to condemn"), contrastare ("to contrast"), debellare ("to defeat"), punire ("to punish"), and sconfiggere ("to defeat").

Concerning the word contraffazione ("counterfeiting"; frequency: 4.422; 0.00022%), the highest level of salience corresponds to the construction contraffazione del marchio ("trademark counterfeiting"; 6.79; 96 tokens); the most salient domains correspond to alimentare ("food"; 14), farmaceutica ("pharmaceutical"; 2). The idea of a partial overlap of pirateria and contraffazione appears to be validated by the and-or pattern, where the two words are associated in 110 tokens. A similar picture comes from the and-or pattern instantiated by contraffatto ("counterfeit") and piratato ("pirated"), which shows a medium value of salience (6.34; the highest level is 10.5 and corresponds to alterato ["altered"]).

A last consideration is deserved by metaphors retrieved from itWac. It is well known that conceptual metaphors offer a rich source for the discursive construction of the world as a function of contextual constraints (Lakoff & Johnson, 1980; see also, for example, the analysis of media coverage of September 11 proposed by Lakoff, 2001), both indexing and constructing culture. This is true to the extent that some of these uses come to be normalized, e.g., included in dictionaries. With their extended meanings, pirate and piracy are metaphors themselves: strictly speaking, such conventional metaphors are “secondary senses” insofar as they activate only certain elements of the meaning potential of the words involved (Hanks, 2006, p. 19). As we have seen, the conceptualization of pirateria (digitale/informatica) is (more than that of contraffazione) associated with the semantic sub-structures of "damage" and "theft" and, at the same time, a pirate is mainly identified as a computer pirate.

Expressions such as pirati di internet ("pirates of the Web"; salience: 8.4), pirati del videogame ("videogame pirates"; 155), pirati del software ("software pirates"; 1.14), pirati del download ("download pirates"; 1.09), pirati della musica ("music pirates"; 0.77) and pirati della rete ("web pirates"; 0.58) show a possible border between piracy "via information technologies" and other forms of piracy. Such border is represented by the difference between the schemas container/containment and path. Indeed, in metaphorical patterns such as pirati della clonazione (referring literally to the cloning of foetuses; salience: 2.15) and pirati della strada (4.89) a (damaging) contact between an agent and a patient is implied while, in the previous expressions, the prominent component of meaning structure is represented by goods (e.g., music, videogames) which enter a possessive relationship with the agent.

The corpus analysis of the word war (a relevant lexeme in the source domain of legal action is war), based on the syntagmatic context noun+preposition+noun, shows that pirateria and pirata always precede counterfeiting in the lists of salience. Guerra contro i pirati ("war against pirates") shows a value of 4.54, while guerra contro i falsari ("war against counterfeiters") shows a value of 3.91. More significantly, guerra al pirata ("war against the pirate") shows a value of 4.56, followed by guerra all'hacker ("war against the hacker"; 4.00), guerra al peer to peer ("war against peer-to-peer"; 3.34), and guerra al falsario ("war against the counterfeiter"; 3.13). Guerra alla pirateria ("war against piracy") shows a value of 5.85, while guerra alla contraffazione ("war against counterfeiting") shows a value of 3.21. In other syntagmatic contexts, pirateria is represented as a “plague” (piaga della pirateria), while the conceptual metaphor of war is also instantiated by constructions such as fronte della pirateria ("front of the piracy").
Conclusions

We have seen that the semantic structure of pirateria, pirata is shaped as a dual conceptualisation (path and container/containment). This result is reflected in the overlap of "piracy," "information technology," and "theft/illegal appropriation." In such a dominant context model, the prominent intention is that of damaging someone or something (e.g., film industry, copyright), while the prominent purpose is that of illegally taking possession of something.

The emerging paradox of this mental model is represented by the fact that (at least from the perspective of piracy culture) there is no "appropriation" or "theft" in downloading and sharing multimedia content, "duplicating" a videogame (as well-known, what is obtained is not a mirror copy of the original product, e.g., DRM systems are not transferred to the copy) or modifying the source code of software. The access to this kind of content can also be perceived (and conceptualized) as a right; such cultural content is not perceived as "stolen" from an owner. For example, it is not by chance that BitTorrent's architecture combats so-called "free riding," which is the term used in P2P networks to describe users who download without sharing (Cardoso, Espanha, Jacobetty, & Lima, 2009). We could also consider strategic choices, such as those related to the game Civilization (which sells millions of copies, despite its "soft" DRM) or to the catalogue of films made available by Warner Bros. through a streaming service, on the social network Facebook; or to the cloud-based methodology which allows "free to play" models of consumption in which only special content is purchased.

This missing perception of the meaning of "appropriation/theft" or "damage" of the "piracy" actions sketched above is mirrored in the concept of "fake." Such concept shows that the traditional dualisms of true/false or material/immaterial cannot be used anymore as organizing principles of goods production and of legal, political decisions and actions concerning communication. This is, obviously, a consequence of technical reproducibility (Benjamin, 2008), which overrules the concept of "authenticity" and generates a quantitative series of events rather than an event (e.g., a film), multiple originals rather than one original. Through complex cultural movements (see the concepts of "mediascape" and "technoscape" in Appadurai, 1990), technical reproducibility (from which the crisis of copyright also stems) seems to bear a new global political system: Western youth cultures, as well as non-Western youth cultures are acting on the same target, that is, the ex- (neo-) colonial power (on this issue, see Bhabha, 1994).

This scenario has been (fore)seen by directors like Roman Polanski, who (along with many others) contributed to a positive, romantic representation of the pirate. It is undeniable that a crucial substructure of the shared meaning of pirate comes from cinema as a popular art. On the ground of the

21 On these issues and the related concept of “loss of profit” see, for example, the following articles:
rebellion and tension toward social change, which characterizes cyberpunk culture (e.g., the power of multinational companies, hackers as the main characters of works such as William Gibson’s short story “Burning Chrome”, 1982), many films have represented the pirate, at least in part, as a positive, romantic character. From a cognitive-linguistic perspective, in view of such a fluidity of the concept “pirate,” an analysis of a large corpus of films could show that rhetorical representations of pirates can be patterned as “conceptual blendings” (Fauconnier & Turner, 2001).

Concerning linguistic data on counterfeiting and piracy, we should note that counterfeiting is invariably a profit-making organization, while “computer piracy”—as it is intended in the analyzed media discourse—is almost always not. Here comes the (language) stigmatization. The projected mental images of piracy may bear no direct resemblance to the real-world phenomena which give rise to them; the problem is, “besides a representation of the meaning of a text, language users also construct mental models of the events texts are about” (van Dijk, 2008, p. 58)—and an ideologically biased representation of events in mental models is what we have seen in language use. A fuzzy legal situation, such as that of digital piracy, appears to receive a relatively negative evaluation when compared to similar (but differently symbolized) phenomena.

As shown by Van Dijk (2008, pp. 192–193), the functions of rhetorical structures may either be semantic (conceptualization: the representation of things in—possibly biased—mental models of events) and pragmatic (felicity: the use of such structures to make discourses adequate to the social situation—in this case, to the relationship with specific groups of interest). The contextual domain of rhetoric is socially shared opinion; the concept of “attack,” “attempt” is not just encoded in the etymology of pirate, but as we have seen, is also reflected in corpora data. Attacks are damaging and, more important, secondary senses which can be classified as metaphors (as pirateria and pirati) “resonate” (Black, 1962) with other words in the immediate context. That is, the salient features of a subject (e.g., “damage” and “theft” for pirateria) guide the reader in interpreting the correlated subjects (e.g., guerra [“war”] in the metaphorical pattern guerra ai pirati). This appears to be the perspective through which it is possible to understand the stigmatization of digital piracy in the light of linguistic data: Since the (dominant) conceptualization of piracy entails an intentional damage as well as theft as a purpose, it is not by chance that conceptual metaphors related to war are used for digital piracy much more than for counterfeiting.

---


23 In the emblematic case of Polanski’s Pirates (1986, starring Walter Matthau as Captain Red), such representation emerges principally from the juxtaposition of the main character and his antagonists (Spanish nobles), who have appropriated an Aztec gold throne. The positive connotation of Captain Red (also supported by the humorous aspects of the character) seems to prevail on his negative features (e.g., violence, potential cannibalism), because the nobles are no less violent and unfair than he is (in this sense, Red is one who steals things from thieves). It is clear that, within an emic perspective, “pirate” gains a positive, romantic valence. Such valence is related, for example, to the success obtained by the PiratPartiet and the Piratenpartei at the elections in Sweden (2009) and Germany (2011).
Further research (e.g., in other languages) on media discourse could show the predictive value of image schemas with respect to context models. At the present time, the use of words such as website and user shows the core of the problem exploded with the court order involving Yahoo! and with the Piracy manifesto “Don’t make me steal” (see #2 in the Appendix). Web users and (the dominant conceptualization of) pirates represent groups of people whose (at least partial) overlap cannot be ignored anymore; such overlap also shows that the analyzed labels are not appropriate to designate the present status of co-creative groups of subjects. This is visible in society (as the hearing of the Antitrust Authority shows), as well as in the symbolization of social and media relationships through language. The present one appears to be a generation which sees cultural products as "free by definition" (see # 2 in the Appendix); for example, concerning films, Cardoso, Espanha, Jacobetty, and Lima (2009) identify such consumers as “innovative” (vs. “traditional” and “mainstream” consumers) within a networked culture—the new media generation group characterized by more intensive weekly habits and watching experiences on multiple screens and multiple platforms (including using the computer to watch and share movies downloaded from the Internet). The concept of “cognitive surplus” (Shirky, 2010) appears to be more than crucial, both in this scenario and with respect to creativity (Pearce, 2011). Perhaps the overlap pirate/Internet user (“Today, every man with a computer is a Producer and a Pirate” states the Piracy Manifesto; see Manetas, 2009) could represent the basis of a possible change in the perception and conceptualization of piracy itself.
References


Appendix – WoP corpus

2. http://blog.sentieriselvaggi.it/?p=1108
3. http://www.lineaedppmi.it/articolo.php?aId=0000089134&cId=43&cpId=7&n=Pirateria+in+azienda%3A+cosa+si+rischia
5. http://www.lineaedppmi.it/articolo.php?aId=0000089134&cId=43&cpId=7&n=Pirateria+in+azienda%3A+cosa+si+rischia
32. http://www.ilvelino.it/articolo.php?id=1320795&t=Web_per_%93About_Elly%94_Yahoo_%EF%80%86_indirizzare_solo_al_sito_ufficiale