A history of Classification and JokeR's Reach

Campen Paul[a], Digue Albin[a]

[a] *Université de Bretagne Occidentale, 20 rue Duquesne, 29200, Brest, France*

Abstract

There have been many attempts to classify humour. Some intended to find permanent features allowing either the concrete creation of humour or a theoretical approach. This essay will serve as a base for people who wish to learn about JokeR's classification reach. Its innate objective is to help with humour translation and in order to do so, it divides the process into these three steps: understanding, translating and recreating wordplay. Those steps go through a stage of development and tests for the following decryption method.

Keywords [1]
Wordplay, classification, linguistic feature, one-liner, single word

1. Introduction

The field of stylistic devices constantly evolves because of its daily use, because of authors playing with language and trying to find new ways in for ambiguity. The free and limitless evolution of language has hindered many scholars and authors trying to establish a stable and permanent classification.

The academic JokeR project [1] strives to automatically classify wordplay to help with its translation. To this end, a new classification was created, tackling this problem mathematically.

We'll go through the previously created classifications to illustrate the subject's evolution and the way people perceived this complex field changed over time.

2. A History of Classification

First of all, it is necessary to quickly summarise how wordplay classification has evolved with time, and what logics were used:

The first created classification, important enough to serve as a beginning for debate, was created in 1976 by Pierre Guiraud [2]. In his work, he tried to exhaustively list every type of wordplay in various categories:

1. Chain wordplay
   a. False coordination: *He works his work, I mine.*
   b. Homophony: *messages that mess ages.*
   c. By echo: *Cool Raoul.*
   d. By habit: *Trois petits chats, chapeau de paille, paillasson, somnambule…*
   e. "Charade à tiroir": *My second is a passionate craftsperson, it is Lovecraft because he Loves craft.*
2. Inclusion wordplay
   a. Anagrams
   b. Palindromes / Anadromes
   c. Spoonerisms

---
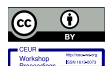
d. Back slang
           e. Acrostics
           f. Acronyms
           g. Interpolation
    3. Substitution wordplay
           a. Homophones / allographs
           b. Homonyms
           c. Paronyms
           d. Synonyms
           e. Antonyms
    4. Puns
           a. Semic
                 i.   by concrete / abstract opposition
                 ii. by polysemy
                 iii. synonymic
                 iv. antonymic
           b. Phonic
                 i.   Homonyms
                 ii. Homophones
                 iii. Paronyms / aphorisms
           c. In absentia
           d. In praesentia
           e. Complex puns

We can recognise the categories for wordplay taking place in unique sentences with the text morphology, and in any kind of context. This classification is complete enough to be kept, but other scholars tried to establish different logics and goals. Here, categories are adapted to the context of the wordplay.

However, one of Freud's [3] wordplay examples states: "Louis XV voulait mettre à l'épreuve l'esprit d'un de ses courtisans, dont on lui avait vanté le talent ; il lui ordonna de faire, à la première occasion, un mot d'esprit sur lui; le roi lui-même, disait-il, voulait lui servir de « sujet »; le courtisan répondit par ce bon mot: « Le roi n'est pas un sujet»." (One day, king Louis XV decided to test one of his courtier's intelligence, who had been presented to him as a man of wit. At the first opportunity, he ordered the courtier to make a witticism with himself, the King, as the subject. The courtier answered, cleverly, "The King isn't a subject". [free translation]). Here, the word "*sujet*" is polysemic for "subject" and "vassal".5

This wordplay shines light on an uncertain boundary between semic puns with a concrete and abstract opposition and semic puns by polysemy.

In 1996, Delabatista [4] suggested a new, simpler classification. Compared to Guiraud's work, we can imagine that Delabatista wanted to simplify classification. This same work directly inspired Gottlieb [5] who, by modifying Delabatista's work, took the question in another direction by detailing the "homonymy" section. With these two scholars' work put together, we get this classification:
    1. Homonymy
           a. Lexical homonymy
           b. Collocational homonymy
           c. Phrasal homonymy
    2. Homophony
    3. Homography
    4. Paronymy

This classification concentrates much more on a single linguistic unit, one word that would be compared with another implied one. This classification is in that way not a simple condensed version of Guiraud's work, as it allows us to tackle the subject differently.

The same year, Leppihalme [6] will divide wordplay possibilities. According to him, wordplay possibilities base themselves on various ambiguities:
    1. Pronunciation

2. Spelling
3. Morphology
4. Vocabulary
5. Syntax

By detailing wordplay integration processes, research allows a step back and a more general view of the subject. Leppihalme's work still has influence today, as visible in the following examples.

By collaborating with Robert French, Jacqueline Henry [7] developed a strict methodology for wordplay translation. In her book, <u>La traduction des jeux de mots,</u> she describes how various people ventured their theory for wordplay classification, which should be used to help with translation. Incidentally, she expanded on Pierre Guiraud's word classification by adding:

1. "Bouts-rimés" (lists of words that rhyme, given to a poet, who has to make a poem with the rhymes in the same order that they were placed upon the list)
2. Surrealist games (exquisite cadaver or automatic writing)

Two years later, in 2005, Yuan Chuandao [8] classified wordplay contextually. Wordplay exists because of the context, a certain way of talking or a specific logic:

1. homonymic wordplay - same sound and writing
2. lexical meaning wordplay - polysemic words
3. understanding wordplay - the innate meaning of a sentence is revealed by its context
4. figurative wordplay - an explicit comparison or metaphor for an implicit figurative meaning
5. logic wordplay - Rhetoric feature with a situational implication depending on the actual context

Giorgadze's classification [9] was written after the study of previous researches.

1. Lexical-semantic wordplay: homonymous or polysemous words
2. Structural-syntactic wordplay, happens when a complex phrase may be understood in different ways. It may use a response to answer the ambiguity
3. Structural-semantic wordplay, depending on the way a sentence is constructed, many meanings can merge and be understood from it

Here, Giorgadze almost took the path of the first parametric classification. We can see 4 different settings working together.

Lexical reach is about the form of units ; the semantic setting is about the meaning carried by units ; the structural setting works for the forms in which the wordplay takes place and acts as context ; the syntactic setting is about how the sentence is constructed to create a wordplay.

3. Objective of JokeR's Classification and Modification

The JokeR project's [1] innate objective is to enable machines to read, decipher and interpret wordplay. To this end, they need a judgement system just as we do, but mathematically constructed.

3.1. Parametric Objective

The base of this parametric classification requires that each wordplay example fit in a single embranchment of precise parameters.

In the cases where an entry meets the conditions for different parameters, we need to teach automatic translation engines which combination is the strongest so that we, humans, can easily perceive wordplay.

With three unique branches, each of them dividing into two to five categories, we can see possibilities multiplying. Our objective is thus to explain how to give a summary of an extremely vast field of study like wordplay with this kind of classification, by analysing each possibility.

The classification was designed to be easy to learn and use. However, work and processing of the data revealed that improvements may be necessary. Thus, some terms may differ from the official version [1].

As an introduction, and to help you understand the following table, we'll explain a few important differences.

"Opposition" is used to define a case where the linguistic unit is opposed to its polysemy or to the sentence's innate meaning.

"Sound", "Writing" and "Both" were reworked to be understandable as the medium through which you may encounter and understand a pun.

The "Other" parameter is different from the preceding media. It comprises wordplay based on syntax or grammar.

3.2. Classification as an Arborescence

Even though this classification works with strict parameters, it is important to know that every single combination is unique and tends to a different type of wordplay.

The following tables are organised in the following way: A strict definition (with an addition in case the wordplay acts on different parameters, in which case the strongest medium has priority), some category of wordplay that fits in with the given parameters, and two examples when possible, one of which is a single-word and the other a phrase-based.

Table 1. List of Vertical possibilities

	Sound	Writing	Both
Identity	Unit that can only have a double meaning orally. The different interpreted meanings have an identical pronunciation but a different writing (can be writing similarity). Homophone, "kakemphaton" (phrasal homophones), holorhymes, pataquès (far-fetched liaisons) Ex: Tadmorv Par les bois du djinn où s'entasse de l'effroi / Parle et bois du gin ou cent tasses de lait froid	Unit that can only have a double meaning in written form. The different interpreted meanings have identical writing but a different pronunciation (can be sound similarity). Homograph Ex: Mes **fils** ont cassé mes **fils**. Cet homme est **fier**, peut-on s'y **fier** ?	Unit with a semantic, not morphological double meaning. The different interpreted meanings have an identical pronunciation and writing. Pun, homonym Ex: « **Haut les cœurs**, mon cher. Le style vient en grande partie de la façon dont vous vous tenez. La poitrine bien haute, comme ceci. » Il entreprend de réaliser un opéra
Similarity	Unit that can have a double meaning because of an oral confusion. We can hear two or more meanings with a similar pronunciation (can be writing identity). Homograph Ex: Larmes à feu	Unit that can have a double meaning because of a written confusion. We can read two or more meanings with similar writing (can be sound identity). Heterograph Homophone Ex: - Il y a du monde à l'intérieur ? - J'en compte peu, **de dents**.	Unit with an oral and written double meaning, with the use of similar words or the modification of a well-known expression. Paronym, portmanteau words close to one of the original words Ex: Se jeter dans la gueule du dragon purrfect
Abbreviation	Oral unit pronounced in a certain way. It can be written normally but will be shortened when spoken. Ex: "- Tu as vu **c't'armoire** ? - Seulement le 7"	Unit with a shortened writing, which will still be pronounced similarly (can be sound identity). Ex: tadmorv wimessir	Unit made out of at least 2 different words, which influences the unit's general comprehensibility. Portmanteau words, acronym Ex: SNCF Cloportuniste
Permutation	Unit where sound inversion creates confusion. If certain changes are made to keep the sound, the wordplay will be exclusively oral. Spoonerism Ex: Plante à **pipa**illons (the original word, "papillon" only has one **I**, but to make the wordplay work with its pronunciation, two **I**s were indispensable) You were **f**ighting a **l**iar in the quadrangle	Unit where the inversion of letters and syllables completely changes the word. The wordplay only works in written form, and isn't understandable orally. Anagram, ananym, palindrome, ambigram Ex: Roma / Amor bestial / bétails / baliste / établis	Unit where sound and syllable inversion changes the general meaning. Can be identifiable orally and in written form. Spoonerism, back slang Ex: **Sonn**ez, les **tromp**ettes ! / **Tromp**ez, les **sonn**ettes ! Une **g**amelle de **m**orilles / une **m**amelle de **g**orilles
Opposition			Unit where the word plays against the innate meaning of a sentence or of itself. Ex: **Énervé** (to be fed up / to be deprived of nerves, thus of reactions) Ce film était **écoeurant** !

The following table has an additional column to mitigate the lack of scope of the original classification. At the earlier stage of this project, we thought that most of the encountered wordplays would play *with* words, but our research revealed various cases which play *on* words. Where other

categories logically define the integration process of wordplay by the audience, the last one includes plurivalent elements.

A few types of these elements are: *chiasmus* which mix polysemic repetitions, *zeugmas* where a given expression gets two meaning in the same sentence, which are often separated with a coordinating conjunction, and *antonymy* which could be perceived as horizontal identity wordplay acting on both sound and writing, but is actually hard to keep to one category.

Table 2. List of Horizontal possibilities

	Sound	Writing	Both	Other
Identity	Unit that can only have a double meaning orally. The different interpreted meanings have an identical pronunciation but a different writing and are identifiable through context (can be writing similarity). Homophone Ex: Un homme **sot** tenait en une main un **sceau** royal et en l'autre un **seau** d'eau. D'un mouvement, les trois **[so]** tombèrent.	Unit that can only have a double meaning in written form. The different interpreted meanings have identical writing but a different pronunciation (can be sound similarity). Homograph Ex: Les poules du **couvent couvent**.	Unit with a semantic, not morphological double meaning. The different interpreted meanings have an identical pronunciation and writing. Polysemous homonym Ex: **Orange** lockers smell like **oranges**!	Unit which plays on different language aspects to create a cryptic linguistic effect. Chiasmus "Janotism" (rupture in syntactic logic) Ex: We **shape** our **buildings**, and our **buildings shape** us.
Similarity	Unit that can have a double meaning because of an oral confusion. We can hear two or more meanings with a similar pronunciation (can be writing identity). Homograph, paronym, assonance, alliteration Ex: Dover andoverandoverandoverandove r	Unit that can have a double meaning because of a written confusion. We can read two or more meanings with similar writing (can be sound identity). Homophone, paronym Ex: Better a **hatter** than a **hater**	Unit with a semantic and morphological double meaning, but with different writing and pronunciation. Paronym, synonym Ex: **Irruption** / **éruption** L'**incident** n'était qu'un **incendie**	Unit which plays on different language aspects to create a cryptic linguistic effect. Distant antonymy (frame opposition) Ex: We have **warm** hearts for **cold** noses
Abbreviation			Unit formed of at least two different words. We can see the source words and the target abbreviation. Portmanteau word, acronym Ex: My opponent was both **powerful** and **vicious**. You might say he was... "**powericious**".	Unit in which repetitions are avoided, thus creating confusion or reducing understandability. Zeugma Ex: "He works **his work**, **I mine**" - *Ulysses*
Permutation	Unit where sound inversion creates another meaning. Is sometimes based on mistakes, on a slip of a tongue rather than on conscious wordplay.	Unit where the inversion of letters and syllables creates a new word. Anagram, ananym, palindrome Ex: "**À révéler mon nom**, **mon nom relèvera**." - Cyrano de Bergerac	Unit where sound and letter inversion change an easily identifiable expression. Spoonerism, Chiasmus Ex: Je vous sal**ie** ma r**ue** / je vous sal**ue** Mar**ie**	
Opposition				Unit where the lexical field contradicts itself. Beyond frames, the two units are clearly antonymous. Antonym, pleonasm Ex: **Find** us to get **lost** | **Monter** sur un podium en **descente**

4. Conclusion

This new classification chooses a much more mathematical approach than the previous ones. With a few adjustments, new types of wordplay could easily be included in this process.

One of the main points is the versatility of such a classification. Does it take into account every listed wordplay entry?

It must be noted that the work began on a non-exhaustive wordplay list which is gradually growing. By discovering a problem with the scope of our classification, a colleague, Julien Boccou, had the idea of introducing an OPPOSITION category. That same reason led us to create an OTHER category.

Our wordplay list has grown with time, but here are a few examples of categories that can't be classified by JokeR yet:

1. Amphigouris, ex: *Vice et versa, Les Inconnus*
2. Univocalics and derivations, ex: No cool monsoons blow soft on Oxford dons
3. Haplologies, ex: urine analysis → urinalysis
4. Lipograms, ex: A crimson bloom of an unknown brand is just as fragrant to an olfactory gland [10](Giordano, s. d.)
5. Compound words, ex: grandmother, darkroom, scarecrow, maidservant
6. Neologisms, ex: Oompa-Loompa

However, like Yuan Chuandao [8] once said, wordplay is bound to evolve along with the world and its languages. This classification strives to install itself and persist. In order to do so, it needs more in-depth study to find a stable solution to this remaining classification question.

5. References

1. Ermakova, L., Miller, T., Regattin, F., Bosser, A.-G., Mathurin, É., Corre, G. L., Hannachi, R., Boccou, J., Digue, A., Damoy, A., & Jeanjean, B. (2022). Overview of JOKER@CLEF 2022 : Automatic Wordplay and Humour Translation Workshop. 25
2. Guiraud, P. (1976). Les jeux de mots. Presses universitaires de France
3. Freud, S. (2019). *Le mot d'esprit et ses rapports avec l'inconscient*. BoD-Books on Demand.
4. Delabastita, D. (1994). Focus on the pun: Wordplay as a special problem in translation studies. Target. International Journal of Translation Studies, 6(2), 223-243
5. Gottlieb, H. (2016). You got the picture?. *Traductio: Essays on punning and translation*, 207.
6. Leppihalme, R. (1997). Culture bumps: An empirical approach to the translation of allusions (Vol. 10). Multilingual Matters
7. Henry, J. (2003). La Traduction Des Jeux De Mots. Presses de la Sorbonne Nouvelle
8. Chuandao, Y. (2014). English Pun and its Classification.
9. Giorgadze, M. (2014). Linguistic features of pun, its typology and classification. European Scientific Journal
10. Giordano, C. (2018). *What Is a Lipogram?* Owlcation. https://owlcation.com/humanities/Word-Play-What-is-a-Lipogram