# Medical term formation in English and Japanese: A study of the suffixes -gram, -graph and -graphy

**Carlos Herrero-Zorita, Clara Molina, Antonio Moreno Sandoval**

Universidad Autónoma de Madrid

This paper presents a translingual study of medical lexicology in English and in Japanese that compares the meaning and usage of three suffixes often found in medical discourse: *-gram*, *-graph* and *-graphy*. By means of an in-depth observation of frequency counts and semantic profiling in actual usage, we present a proposal regarding which roots each of the suffixes allow, together with an analysis of the meaning subtleties of the affixes. This work, informed by both cognitive and corpus linguistics, advances the presence of a concurrent pattern in English-Japanese morphology within medical discourse. After presenting a number of significant parallelisms and differences between both languages within the corpora, however, the work concludes with an explanation of how and why the three suffixes under inspection display quite distinct meaning nuances that restrains them from being used at random, both in English and in Japanese.

**Keywords:** medical terminology, derivative morphology, Japanese, English, conceptual blending.

## 1. Introduction

This work compares the way in which two languages, English and Japanese, make use of three highly frequent suffixes in medical discourse, namely, *-gram*, *-graph* and *-graphy*. Medical discourse stands as a rich data source for lexical enquiry, if only for the large number of long-standing studies on the matter, ranging from the Hippocratic corpus to Harris's theory of the sublanguage at the basis of Natural Language Processing (Friedman, Kra & Rzhetsky, 2002), to name but a couple. By means of combining cognitive and corpus linguistics, we aim at, on the one hand, contributing to the field of medical terminology and, on the other, further stressing the natural rapport between the fields of cognitive and corpus linguistics (as attested in, among others, Hollmann, 2006; Fischer, 2010; Fuertes-Olivera & de Alba, 2011;

Ibarretxe, 2010; Soler, 2008; or Wulff, 2006), in that they both adopt a usage-based analytical approach to actual patterns in natural contexts. To the best of our knowledge there are no results in the literature regarding medical morphology in Japanese and English, though we should take into account similar non-English lexicological studies such as Ibarretxe-Antuñano, 2006; Jiang, 2012; and Türker, 2013.

The data for this study comes from two specialised corpora, the Japanese medical corpus from the MultiMedica project carried out at the Computational Linguistics Lab at Universidad Autonoma de Madrid (Moreno-Sandoval & Campillos-Llanos 2013), and the scientific subcorpus of the British National Corpus. As advanced above, the study focuses on the suffixes *-gram*, *-graph* and *-graphy*, and it consists of an analysis (i) of each of them, and (ii) of the roots they get attached to, both in English and in Japanese, at all times bearing in mind both meaning nuances and usage frequencies as displayed by real data.

## 2. Case study

Word formation, although a universal process, is very different in Japanese and English, especially regarding medical term formation. English medical terms are formed by affixation, adding Graeco-Latin prefixes and suffixes that have changed very little from their original forms (Creason, 2010, pp. 7). For example:

(1) *Cardiogram* ('a record of muscle activity within the heart made by a cardiograph') is constructed with the root *cardio-* ('heart', from Latin *cardi-*, which originally evolved from Greek *kardia*) and *-gram* ('denoting something written or recorded').

(2) *Magnetograph* ('a device for detecting and recording variations in the intensity and direction of magnetic fields') is formed by *magneto-* ('magnetic' or 'magnetism', from Greek magneto) and *-graph* ('instrument for recording; something written').

(3) *Phlebography* ('The radiographic examination of veins injected with a radiopaque contrast medium') is constructed by *phlebo-* ('vein' from Greek *phleps*) and *-graphy* ('process of writing or recording').

Japanese, however, does not have Graeco-Latin origins, and medical terms are not formed by affixation. Many Japanese medical terms have their origins in ancient Chinese medicine. The Chinese terms were adapted to Japanese kanji[1] around the 9th century (Izumi & Isozumi, 1990, p. 91). The majority of the terms used today, however, are borrowed from Western languages. The acquisition process began through the medicine books traded with Dutch merchants during the almost complete closure of Japan in the 17th and 18th centuries, and then more rapidly with the adoption of the German medical system in the 19th century (Irwin, 2011, p. 37). The medical terms were introduced in Japanese by means of two different processes: on the one hand, the translation and coining into Sino-Japanese compounds using kanji. Since Japanese is an agglutinative language, the majority of medical terms written in kanji are formed by composition using free morphemes, instead of affixation as in English (Herrero-Zorita, 2013) (see examples 1-3). When a medical term is created, and since Japanese is a right-headed language (Miyaoka & Tamaoka, 2005, p. 46), the free morphemes that correspond to our English suffixes are attached to the right of the root: 図 'Diagram' attaches to 聴力 'Hearing ability' and creates 聴力図 'A diagram of the hearing ability' (An *audiogram*). Hence, in a similar way to English suffixes, the root receives the medical morpheme, which defines the meaning of the complete word (see example 4). The second possible process of term formation would be the direct transcription into the katakana alphabet (see example 5).

(4) Composition of the word *arteriography* in Japanese:

動脈 + 造影 = 動脈造影

artery + imaging = 'imaging of an artery'

arterio- + -graphy = arteriography

[1] Japanese combines three writing systems: kanji (ideograms of Chinese origin), hiragana (a syllabic system from Japan) and katakana (also a syllabic alphabet, mainly used for transcribing foreign words). A fourth, non-Japanese alphabet is used in some occasions, named romaji, that uses Latin characters.

(5) Transcription of the word *chromatogram*:

クロマトグラム

/ku-ro-ma-to-gu-ra-mu/

chromatogram

Therefore, our object of study will be, first, the three Graeco-Latin suffixes used in English *-gram*, *-graph* and *-graphy* and their Japanese counterparts in kanji and katakana.

## 3. Methods

The following steps were taken for the development of the study is simple: first, we examined the definitions of each morpheme, starting from the English ones and then moving to the Japanese counterparts; secondly, we observed their frequency of usage in both corpora; thirdly, we analysed the number of roots to which each morpheme can be added and studied their frequency and meanings.

The examination of the definitions required two types of dictionaries: general and specialised. For the English language we used the *New Oxford American Dictionary* (2005), *Dorland's Medical Dictionary for Health Consumers* (2007), *Stedman's American Heritage Medical Dictionary* (2007), and *Mosby's Medical Dictionary* (2009). For Japanese we used the *Online Life Science Dictionary*[2], developed as part of the *Life Science Dictionary Project* at Kyoto University. For English and Japanese etymological information, we used the *Online Etymology Dictionary* (Harper, 2001) and the *Free Online Kanji Etymology Dictionary* (Howell and Morimoto, 2004).

We used two major corpora for each language. On the one hand, the British National Corpus (BNC)[3] for the extraction of frequencies and examples from the English language. The BNC is a collection of 112,181,015 tokens of written and spoken English from many different discourses, newspapers, academic essays, informal conversations,

[2] http://lsd.pharm.kyoto-u.ac.jp/en/service/weblsd/index.html
[3] http://www.natcorp.ox.ac.uk/

etc. For the development of this study, we have used the subcorpus of natural and pure sciences, consisting of 3,821,902 tokens.

For Japanese, we have used the Japanese corpus from the MultiMedica project[4] at the Computational Linguistics Laboratory of the Autonomous University of Madrid (LLI-UAM). The corpus consists of texts from several specialised journals: *Kampo Medicine* (Japanese oriental medicine), *Kansenshogaku Zasshi* (infectious diseases magazine), *Kanzo* (magazine about diseases of the liver), *ORLTokyo* (Japanese otorhinolaryngology) and *Sanfujinka no shinpo* (advances on obstetrics magazine). The corpus is formed by a total number of 3,746 documents and 997,784 tokens. The following section will provide the definitions of each suffix as well as their frequencies.

## 4. Results

### 4.1. Definitions of the morphemes studied

Regarding the different meanings of the English suffixes, we can see a major difference between the definitions from the general English dictionary and the specialised dictionaries: First of all, the general definitions were quite different between each other. For example, *-gram* may refer to 'denoting something written or recorded', to indicate 'a novelty greeting or message as a humorous or embarrassing surprise', but also to the metric unit of on thousandth of a kilogram (Table 1). The medical definitions, on the other hand, are fewer and more similar between each other, which match to the definitions retrieved from the etymological dictionary: *-gram* referred to the representation of an object, *-graph* to the instrument that produces this representation, and *-graphy* the process or method or representing it. Moreover, *Mosby's Medical Dictionary* also stated that *-graph* is the "product of drawing or writing" (See Table 1).

[4] http://www.lllf.uam.es/ESP/Multimed.html

| Suffix | General Definitions | Medical Definitions | Etymological Information |
|---|---|---|---|
| ***-gram*** | In nouns denoting something written or recorded (esp. in a certain way)<br><br>In nouns denoting a novelty greeting or message as a humorous or embarrassing surprise for the recipient<br><br>Referring to the metric unit of mass equal to one thousandth of a kilogram | *Dorland's Medical Dictionary*: 'written'; 'recorded'.<br><br>*American Heritage Medical Dictionary*: 'something written or drawn'; 'a record'.<br><br>*Mosby's Medical Dictionary:* suffix meaning a 'drawing' or a 'written record'. | From the Greek *graphien* ('to carve', 'to write', into *gramma* ('something written or drawn') to *gram.* |
| ***-graph*** | In nouns denoting something written or drawn in a specified way<br><br>In nouns denoting an instrument that records | *Dorland's Medical Dictionary*: 'a writing or recording instrument'; 'the record made by such an instrument'.<br><br>*American Heritage Medical Dictionary*: 'an instrument for writing, drawing or recording'.<br><br>*Mosby's Medical Dictionary*: suffix meaning 'the product of drawing or writing'; 'a machine for making something drawn'. | From the Greek *graphien* ('to carve', 'to write', into *grapho* ('to write'), then specialised into the instrument with *graph.* |
| ***-graphy*** | A descriptive science<br><br>A technique of producing images<br><br>A style or method of writing or drawing | *Dorland's Medical Dictionary*: 'a method of recording'; 'writing or recording'.<br><br>*American Heritage Medical Dictionary*: 'a writing or representation produced in a specified manner or by a specified process'.<br><br>*Mosby's Medical Dictionary*: suffix meaning 'a kind of printing or process of recording' | From the Greek *graphien* ('to carve', to write', into *graphia* ('description of') to *graphy.* |

Table 1: Definitions provided by general English, medical, and etymological dictionaries

Continuing with the Japanese counterparts, Tables 2 to 4 present the definitions of the Japanese equivalents to *-gram*, *-graph* and *-graphy*. Two observations can be drawn: first, that the three morphemes are used in katakana, which is the phonetic transcription of the English suffixes. Second, that there are several possibilities to each suffix in Japanese, and these relate to similar concepts as the English suffixes: those kanji equivalent to *-gram* denote images, figures and diagrams; the ones equivalent to *-graphy* are related to processes, such as 'imaging', 'exposure', 'photographing', 'examination', 'scan', etc. Between those corresponding to *-graph*, which denote 'instruments', 'devices' or equipment', there is one that can be translated as 'picture' or 'photo', which can also be used for an equivalent to *-gram* (Table 2).

| **Japanese equivalents to *-gram*** | **Definition** |
|---|---|
| グラム | Transcription of Western *-gram* in *katakana* |
| 図 | Figure, plot, diagram |
| 造影図 | An imaged figure |
| 図表 | Chart, diagram, graph |
| 像 | Image, picture, graph |
| 写真 | Picture, photo |

Table 2: Japanese counterparts for the suffix *-gram* and their definitions

| **Japanese equivalents to *-graph*** | **Definition** |
|---|---|
| グラフ | Transcription of Western *-graph* in *katakana* |
| 記録計 | Record measuring device |
| 計 | Measuring device |
| 機器 | Device, equipment |
| 装置 | Equipment, apparatus |
| 写真 | Picture, photo |

Table 3: Japanese counterparts for the suffix *-graph* and their definitions

| **Japanese equivalents to *-graphy*** | **Definition** |
|---|---|
| グラフィー | Transcription of western *-graphy* in *katakana* |
| 造影 | Imaging |
| 検査 | Examination, scan |
| 図検査 | The examination of an image (‘Figure, plot, diagram’ + ‘Examination, scan) |
| 造影検査 | Imaging + Examination, scan |
| 撮影 | Exposure, photographing |

Table 4: Japanese counterparts for the suffix *-graphy* and their definitions

Hence, the definitions of the medical morphemes show similarities between English and Japanese. In both languages *-gram* is related to the representation of an

object, *-graph* to the instrument that produces this representation, and *-graphy* the process or method or representing it. Also in both languages *-graph* is a suffix that appears to contain additional meaning: *Dorland's Medical Dictionary* states as a second definition that it can imply 'the record made by such an instrument', and the Japanese counterpart 写真 can be translated as 'picture' or 'photo' (Table 3), which can also used as an equivalent to *-gram* (Table 2). In order to continue exploring this assumption, we will observe their frequency of usage, explained in the next section.

## 4.2. Frequency of the suffixes in English and Japanese

In the same way as the definitions, the distribution of the frequency of usage of each suffix appears to be similar in both languages. Table 5 shows the frequency of each suffix and also the number of different roots to which they are attached in our corpora. The saliency of the suffix *-graphy* is remarkable in both languages, both in terms of frequency and number of roots to which it is attached. On the other hand, *-graph* appears to be the least used suffix in both languages.

| | **Raw Frequency** | | **Total Frequency Normalised per Million** | | **Number of Roots** | |
|---|---|---|---|---|---|---|
| | English | Japanese | English | Japanese | English | Japanese |
| ***-gram*** | 125 | 79 | 32.7 | 79.2 | 20 | 12 |
| ***-graph*** | 97 | 30 | 25.4 | 30.1 | 9 | 5 |
| ***-graphy*** | 417 | 353 | 109.1 | 353.8 | 23 | 15 |

Table 5: Frequencies of each suffix in the corpora

The frequencies extracted from the corpora allow us to observe how each suffix is used in real language. Once again, both languages behave similarly, having *-graphy* as the preferred suffix, up to four times more frequent than the second most used, *-gram*. Suffix *-graph*, which, as we saw, had a more irregular definition, is in both languages the least used, with a frequency of 25.4 per million in English, and attached to only 5

different roots in Japanese. Our next step in the study will examine these roots that accept each suffix: which are they, and the terms that are formed.

### 4.3. Terms formed by the suffixes and their frequency of use

The analysis of the definitions of the terms that contain the suffixes shows the same pattern as before: the terms containing *-gram* are related to the representation of their root; those that end in *-graph* refer to the instrument referred by the root; and those containing *-graphy*, to the process of representing the root (see Appendix A for details frequencies and definitions). There are, however, a couple of exceptions once again with suffix *-graph*: *radiograph* (and also *autoradiogram*/*graph*) and *photomicrograph* referred to the representation and not the instrument.

In relation to Japanese, the results show concurrence with regard to English: the definitions of the Japanese terms carry the concepts of 'representation', 'instrument', and 'process, respectively, except for the term *radiograph* (X線写真) (See Appendix B). As seen in Table 3, the kanji referring to *-graph* that denotes the concept of 'picture' or 'photo' (写真) is only used for the term *radiograph* and its variants, precisely the one that appears to be ambiguous in English. In fact, dictionaries account as a synonym of the term *radiograph* the term roentgenogram and roentgenograph. The ambiguity of the suffix *-graph* is still present, in both languages. The following and final section will analyse the acceptability of each suffix.

### 4.4. Comparison of the level of acceptability of the suffixes

The following tables present the medical terms that contain our three suffixes with the roots found in the corpora. Each row corresponds to one root, and each column to the suffix. In order to make sure if the term exists –the fact that it does not appear in the corpora does not mean it would not exist in other text– was verified in a four way process: (1) if it appears in the corpus, we accept the term immediately; (2) if it does not appear in the corpus, we verify if it appears in a medical dictionary; (3) if it does not appear neither in the corpus nor the dictionary, we check if it appears in a real medical text from journals found in Google Scholar; (4) if it does not appear in neither of the

three previous cases, the term is discarded. Each case has been indicated in a different pattern:

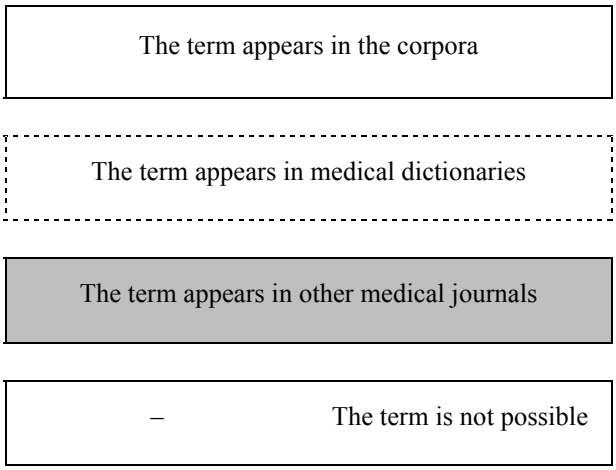


Figure 1: Legend for tables 6 and 8

Table 6 shows the results from the terms in English. The results indicate that -*graphy* allows the highest number of roots, since it is able to appear with nearly any one of them. Suffix -*graph*, on the other hand, shows the opposite situation, where the number of roots to which it can be attached is much lower.

| ***–gram*** | ***–graph*** | ***–graphy*** |
|---|---|---|
| Engram | – | – |
| – | Radiomicrograph | – |
| Angiogram | Angiograph | Angiography |
| Arteriogram | – | Arteriography |
| Autoradiogram | *Autoradiograph: image, not instrument | Autoradiography |
| Barogram | Barograph | Barography |
| Cardiotocogram | Cardiotocography | Cardiotocography |
| Chromatogram | Chromatograph | Chromatography |
| Echocardiogram | Echocardiograph | Echocardiography |

| | | |
|---|---|---|
| Electrocardiogram | Electrocardiograph | Electrocardiography |
| Electroencephalogram | Electroencephalograph | Electroencephalograph |
| Electromyogram | Electromyograph | Electromyography |
| Electropherogram | – | Electropherography |
| Electroretinogram | Electroretinograph | Electroretinography |
| – | – | Fluorography |
| Histogram | – | Histography |
| Idiogram | – | Idiography |
| Immunoscintigram | – | Immunoscintigraphy |
| Interferogram | Interferograph | Interferography |
| Magnetogram | Magnetograph | Magnetography |
| Mammogram | – | Mammography |
| – | Micrograph | Micrography |
| Oscillogram | Oscillograph | Oscillography |
| Phlebogram | Phleboraph | Phlebography |
| – | *Photomicrograph: image, not instrument | Photomicrography |
| Plethysmogram | Plethysmograph | Plethysmography |
| Pyelogram | – | Pyelography |
| Radiogram | *Radiograph: image, not instrument | Radiography |
| Spectrogram | Spectrograph | Spectrography |
| Spirogram | Spirograph | Spirography |
| Stratigram | – | Stratigraphy |
| Tomogram | Tomograph | Tomography |

| | | |
|---|---|---|
| Topogram | – | Topography |
| Tympanogram | Tympanograph | Tympanography |
| Ultrasonogram | Ultrasonograph | Ultrasonography |
| Urogram | Urograph | Urography |
| Ventriculogram | – | Ventriculography |

Table 6: Appearances of the medical terms with each suffix

The results have been quantified and are shown in Table 7. From all the possible roots found in the corpora, a total of 37, *-graphy* can be attached to 35 (94.6%) followed by *-gram* with 33 roots (89.2%) and finally *-graph* with 22 (59.5%).

| **Total number of Roots: 37** | ***-gram*** | ***-graph*** | ***-graphy*** |
|---|---|---|---|
| The term appears in the corpora | 20 | 9 | 23 |
| The term appears in medical dictionaries | 9 | 11 | 9 |
| The term appears in other medical journals | 4 | 2 | 3 |
| Totals | 33 | 22 | 35 |

Table 7: Information on the flexibility of the English suffixes with the roots

Table 8 shows the same operation in Japanese. The outcome is similar to the one in English: the morphemes corresponding to *-graphy* are the most frequent and productive, as they accept more lemmas, and *-graph* is the least productive.

| Japanese *–gram* | Japanese *–graph* | Japanese *–graphy* |
|---|---|---|
| 血管造影図 | – | カラードプラアンギオグラフィー，血管造影，血管造影検査，血管撮影 |
| 動脈造影図 | – | 動脈造影 |
| オージオグラム，聴力図 | – | – |
| オートラジオグラム | オートラジオグラフ | オートラジオグラフィー |
| クロマトグラム | クロマトグラフ | クロマトグラフィー |
| エコー像 | 音波検査器 | エコー検査 |
| 心エコー図，心エコー像 | 心エコー検査装置 | 心エコー検査 |
| 蝸電図 | – | 蝸電図 |
| 筋電図 | 筋電計 | 筋電図検査 |
| ヒストグラム | – | – |
| イムノクロマトグラム | イムノクロマトグラフ | イムノクロマトグラフィー |
| – | マイクロライノグラフ | – |
| – | – | 睡眠ポリグラフ検査 |
| – | – | 門脈造影 |
| – | – | 重心動揺検査 |
| – | レーダーグラフ | – |
| X線写真，X線像 | *X線写真: image, not instrument | X線検査 |
| ガリウムシンチグラム | – | シンチグラフィー |
| サーモグラム | 温度記録計 | サーモグラフィー |
| 断層像 | X線断層撮影装置 | 断層撮影 |
| 超音波像 | 超音波機器 | 超音波検査 |

Table 8: Appearances of the medical terms with each morpheme

As shown in Table 9, from a total of 21 roots, the Japanese *-graphy* can be attached to 17 of them (81.0%), *-gram* to 16 (76.2%) and *-graph* to 10 of them (47.6%).

| **Total number of Roots: 21** | ***-gram*** | ***-graph*** | ***-graphy*** |
|---|---|---|---|
| The term appears in the corpora | 12 | 3 | 15 |
| The term appears in medical dictionaries | 3 | 3 | 2 |
| The term appears in other medical journals | 1 | 4 | 0 |
| Totals | 16 | 10 | 17 |

Table 9: Information on the flexibility of the Japanese morphemes with the roots

All in all, throughout the definitions of the suffixes, their frequencies, and the roots to which they can be attached, the results have shown a similar pattern between English and Japanese: each suffix has a clear definition and pattern of usage. Also, suffix *-graphy* appears to be controversial in both languages, as its definition blends with *-gram*, and its usage is not so extended.

## 5. Discussion of the results

One of the main goals of this research is to observe the meaning and usage of three medical suffixes in English and their Japanese counterparts, in an attempt to find significant patterns. That is, if two languages which are, apparently, very different, may *conceptualise* these medical meanings in the same way and if there is a connection with the written morphemes (Jackendoff, 2002), using at all times medical corpora and dictionaries as evidence. As a result of observing the definitions and frequency use of the suffixes and the roots that allow them, we can outline a series of deductions.

First of all, we can organise the meaning of the suffixes separating the definitions from the general dictionaries from the medical ones. The general dictionaries include the medical meanings, but also offer additional definitions, whereas the specialised dictionaries offer a unique definition for each suffix: *-gram* is related to the representation of an element, *-graph* denotes the instrument that makes this representation and *-graphy* the process of representing it. Therefore, as our starting point, we consider the medical definition as the prototypical definition for our study,

becoming the centre of the different meanings that can appear in general English. The prototypical category is *experiential*, that is, it is linked to the context where it is used Geeraerts, 2006, p. 146). Presenting the different definitions of these suffixes on prototypical grounds can provide the users a clear view on how they behave both today and over time, as well as to see which are and have historically been the most salient ones (Molina, 2008: 20); a saliency that depends on the type of user and the situation, in this case, medical discourse.

For this reason, each suffix is polysemic. The definitions are organised semasiologically, giving preference to their meanings, and can be represented as the following (Figures 2 to 4, adapted from Geeraerts, 2006, p. 152). The medical definition appears in the centre, and the additional meanings surrounding it in layers, the ones closer to the centre sharing more semantic elements than the external ones. The semantic elements are defined by two areas: the etymological origin *graphien* (number 1) and the meanings related to the 'act of writing' (number 2). Taking as an example the *-graphy* suffix (Figure 2), the three definitions share the same etymological origin, but the definition of 'a descriptive science' does not share the concept of 'act of writing'. In the case of *-graph* (Figure 3), as we have seen, the boundaries of its two definitions, the 'instrument of the recording' as the main one, and the 'representation' as the second, are much more unclear.

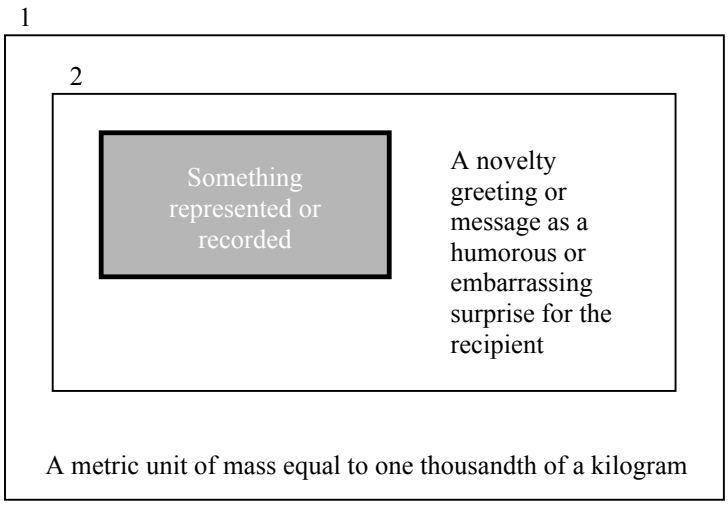


Figure 2: Definitional analysis of *–gram*

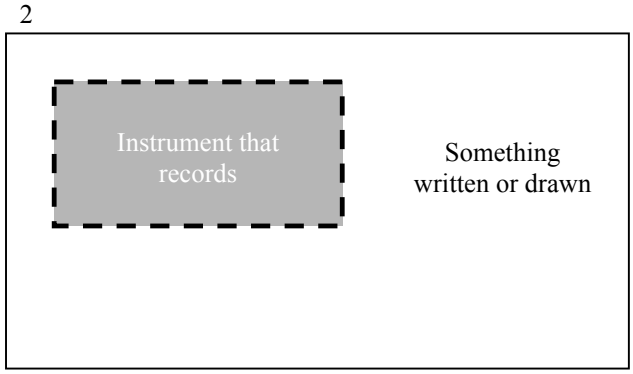


Figure 3: Definitional analysis of *–graph*

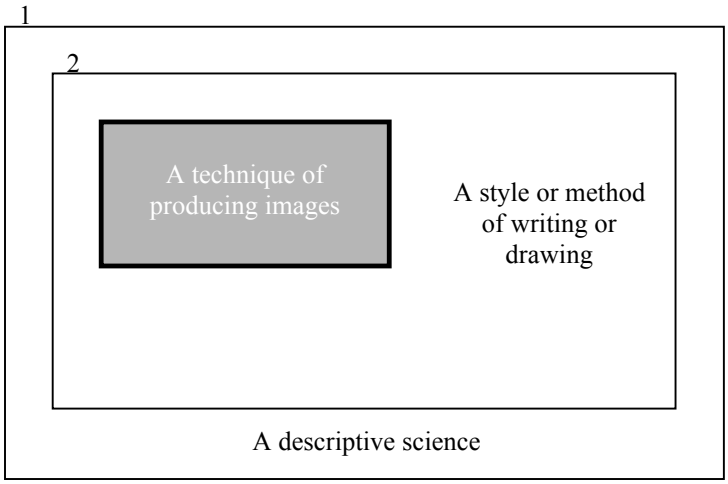


Figure 4: Definitional analysis of *–graphy*

Secondly, focusing only the medical meanings of the suffixes, there are numerous similarities between both languages. In English and in Japanese, *-gram* is related to the representation of the root to which it is attached, *-graph* denotes the instrument that makes this representation and *-graphy* the process of representing it. However, as we have seen, *-graph* appears to have a couple of exceptions, which are also present in Japanese. It can also mean 'image' or 'picture', becoming closely related to suffix *-gram*. Regarding the terms, *radiograph* (and its variants) is the only one that can refer to the representation and not the instrument. The same happens in Japanese with the morpheme 写真 ('picture', 'photo'), which can be the equivalent to both *-graph* and *-gram*, and the term corresponding to *radiograph* (X線写真), which is as in English an image and not an instrument. Recalling Goldberg's view on conceptual formations, "cross-linguistic generalisations are explained by appeal to general cognitive constraints

together with the functions of the constructions involved" (2007, p. 596). However, in terms of medical morphemes between English and Japanese, the historical-cultural point of view may contribute to the explanation: the actual medical practice of Japan began in the 19th century through large influences coming from the West, so it is understandable that the affixes were translated into Japanese carrying the same meaning.

Thirdly, the results indicate that the combination root-suffix in both languages is not constructed randomly. The process of medical term formation, which appears to be similar in both languages, can be represented following the schemas proposed by Bergen and Chang (2007, p. 604), showing the conceptual constructions and the connections between the forms and the function. We have a series of roots with a specific form –either Latin or kanji/katakana alphabet– and meaning, which may or may not accept the three suffixes studied. If the suffix can be attached to the root, the medical term is formed, and its meaning will be defined by the suffix: if it ends in -*gram* it will imply a representation; if it is -*graph* it will be an instrument (or in some cases also a representation); and if it is -graphy it will be a process. Figures 5 to 7 represent our interpretation of these schemas applied to our three morphemes:

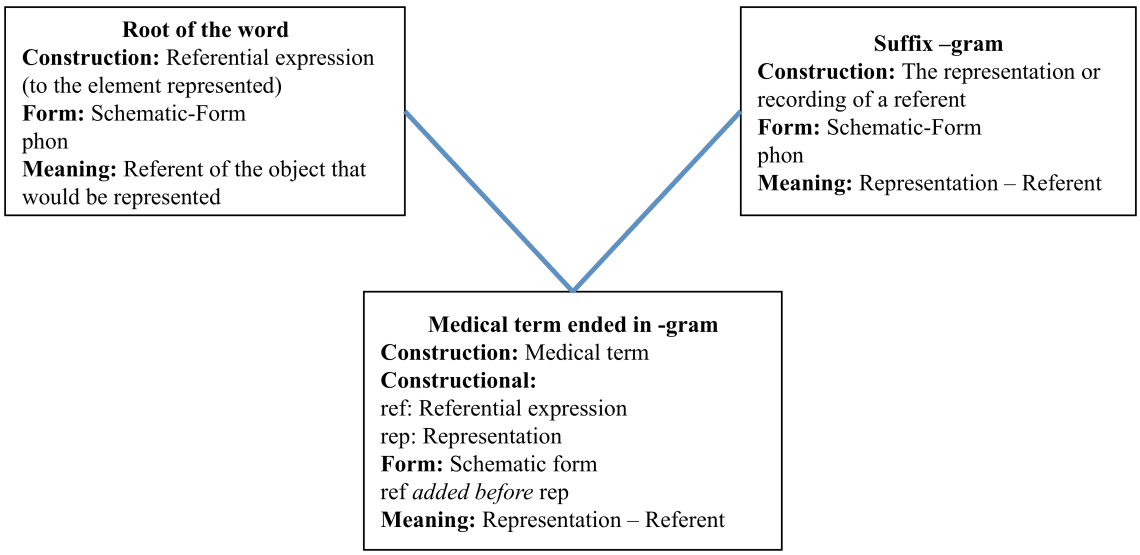


Figure 5: Schematic representation of the suffix -*gram* and its Japanese counterparts

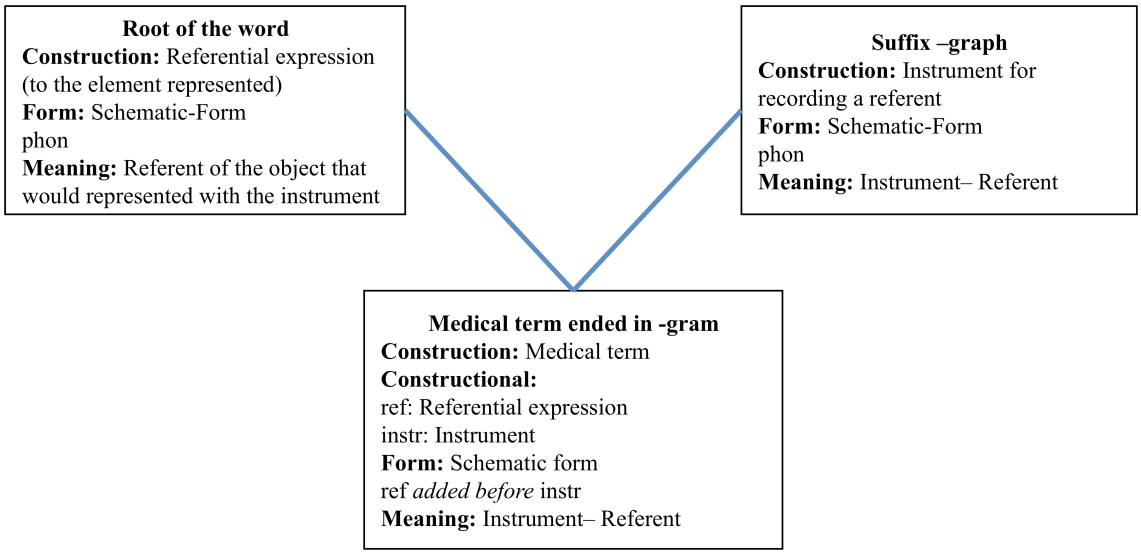


Figure 6: Schematic representation of the suffix *-graph* and its Japanese counterparts

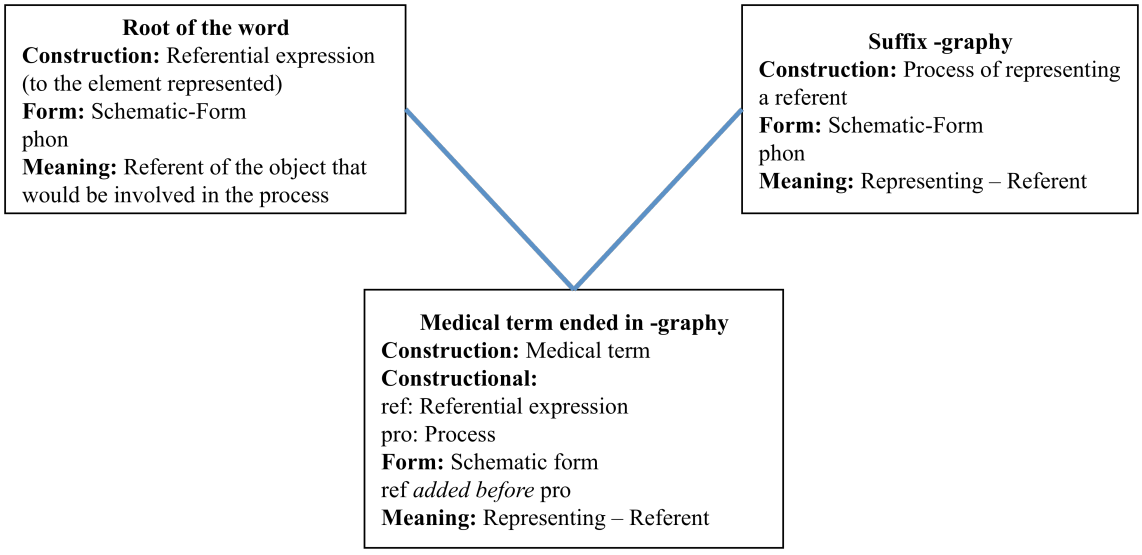


Figure 7: Schematic representation of the suffix *-graphy* and its Japanese counterparts

The main difference between these representations is that in Japanese, the *form* of the morphemes is mostly made through the representation of kanji. We can argue that the union between form and meaning in kanji characters is even tighter than the one we can find in a Latin alphabet, such as in English. Being ideograms –written characters that symbolise an idea without indicating the sounds used to pronounce it– there is an emphasis on form. Moreover, the origins of these characters were actual pictures, representing the objects that were seen by the people writing them. For example, one of the kanji used to represent *-gram*, 図 ("figure", "plot", "diagram"), has the following etymological information:

(6) Formerly 圖, 啚 combines a variant of 嗇, a representation of 'a granary' + 'a square' suggesting 'a particular place', meaning altogether 'a (tall) granary in a specific place'. The final 圖 adds 口, a 'circular enclosure', forming an 'enclosed granary'. From this point the kanji evolved and simplified its strokes becoming 図, and 'diagram' and 'draw' turned into extended meanings interpreted as 'to enclose within the boundaries of a writing medium' (Howell & Morimoto, 2006).

In terms of our three medical suffixes, even though English and Japanese lexically represent the terms in a very different way, there are many similarities in terms of semantic information. The form of each language, using different alphabets, appears to be the most significant difference, but the internal process of medical term formation is equal. It is a shared operation of combination of concepts, in our case, an element that is being represented, an instrument that performs an action, and a process of creating this representation. Based on Turner and Fauconnier's steps in conceptual combination, i.e., composition, completion and elaboration (2002, in Ungerer, 2007, p. 656), the process of creating medical terms in both languages is the following:

- Composition: the union of one suffix to the root. It is a much more flexible process in the case of *-graphy* in both languages, as it can be attached to nearly any root.

- Completion: the addition of background knowledge to the word. Making use of these morphemes to create new medical terms implies using the historical knowledge

attached to them: morphemes with Graeco-Latin origin, and Japanese morphemes that have undergone a process of adoption from languages such as Dutch or English. Bearing in mind this information, for example, a word ended in *-graph* in any medical context will very likely be referring to an object, a tool, which carries out an action, and the same can be applied to Japanese.

- Elaboration: the completion of the concept, passing the tests of correctness and consistency. If a new word is created using these morphemes, there are a series of restrictions in the process, including, for example, the correct spelling of words. Another consideration that has to be made is that not all the suffixes can be attached to all the roots, as we have seen in Tables 7 and 9, where, for example, *-graph* can only be attached to 22 of 37 roots in English, and to 10 from 21 roots in English. All in all, the process is not random.

## 6. Conclusions and further research

The analysis of the three medical suffixes *-gram*, *-graph* and *-graphy* in English and Japanese reveals that there is a parallelism between both languages, and that their usage is not random. Each English suffix, and the Japanese corresponding morpheme, has a salient meaning in the medical discourse, that adds to the root and prevails when the final word is formed: *-gram* denotes a representation of an element, *-graph* to the instrument that performs this representation, and *-graphy* to the process or method of representing it. Also, each suffix is restricted to a series of roots in both languages. Apparently both languages may be very different, especially in terms of form –the alphabets are completely different– and morphology –English uses bound morphemes (suffixes) and Japanese free morphemes. However, the operation underlying the creation of medical terms is very similar. There is an equal combination of semantic meanings. Studying the definitions of our three suffixes, observing and quantifying real examples using corpora, and understanding the process of creating medical terms, can make us understand how the suffixes behave, what meaning they add to the medical term, and predict the formation of new terms, a process that is very common in medical discourse.

Therefore, our study has shown how two languages with dissimilar morphological typology and writing system can behave analogously in medical terminology. The semantic information should be more important than the phonetic and morphological characteristics. We are aware that both corpora used in the study may not be wide enough in order to extract a more solid conclusion; however, the methodology followed could provide support for medical terminology applications such as automatic translation or dictionary creation. It has already been applied for the creation of an automatic term extractor, since the knowledge of how medical terms are formed is crucial for this task (Herrero-Zorita et al., 2014). Finally, the lack of research in comparative studies between English and Japanese medical discourse proves to be an area where there is still plenty of work to do and, given that there are specialised corpora such as the ones used for this study at the disposal of any researcher, there is enough empirical data to perform reliable and useful studies in the future.

## 7. Acknowledgements

This research has been funded by the MINECO (under the grant TIN2010-20644-C03-03) and by the Madrid Regional Government (grant MA2VICMR).

## 8. Bibliography

Bergen, B. K. and Chang, N. (2007). Embodied construction grammar in simulation-based language understanding. In Evans, V., Bergen, B. and Jörg Zinken (eds.) *The Cognitive Linguistics Reader*. London: Equinox Publishing.

Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Creason, C. (ed.) (2010). *Stedman's Medical Terminology*. Baltimore: Lippincott Williams & Wilkins.

Fischer, K. (2010). Accounting for the role of situation in language use in a Cognitive Semantic representation of sentence mood. In *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches* (pp. 101–125). Berlin/New York: Walter de Gryuter.

Gibert (Eds.), *Estudios de fololog*ía *inglesa: Homenaje a la Dra. Asunción Alba Pelayo* (pp. 131–144). Madrid: UNED.

Fuertes-Olivera, P. A., & de Alba, B. P. C. (2011). A Corpus Analysis of Prototypical Causation in Written Scientific and Technical English. *Revista Española de Lingüística Aplicada*, (24), 73–94.

Geeraerts, D. (2006) *Prototype theory*. In Geeraerts, D. (ed.) *Cognitive Linguistics: Basic Readings*. New York: Mouton de Gruyter.

Geeraerts, D. and Cuyckens, H. (2007) Introducing Cognitive Linguistics. In Geeraerts, D. and Cuyckens, H. (eds.) *The Oxford Handbook of Cognitive Linguistics.* Oxford: Oxford University Press.

Goldberg, A. E. (2007). Conctructions: a new theoretical approach to language. In Evans, V., Bergen, B. and Jörg Zinken (eds.) The Cognitive Linguistics Reader. London: Equinox Publishing.

Grondelaers, S., Speelman, D. and Geeraerts, D. (2007). In Geeraerts, D. and Cuyckens, H. (eds.) *The Oxford Handbook of Cognitive Linguistics.* Oxford: Oxford University Press.

Hamawand, Z. (2007). Suffixal rivalry in adjective formation; a cognitive-corpus analysis. *Reference and Research Book News*, *22*(3). Retrieved from http://search.proquest.com/docview/199718333

Hamawand, Z. (2011). Prefixes of Degree in English: A Cognitive-Corpus Analysis. *Open Journal of Modern Linguistics*, *01*(02), 13–23. doi:10.4236/ojml.2011.12003

Harper, D. (2001). *Online Etymology Dictionary*. http://www.etymonline.com/index. php. Accessed 16/02/2014.

Herrero-Zorita, C. (2013). An initial approach on medical term formation in Japanese through the usage of corpora. *Proceedings of the 7th Corpus Linguistics Conference 2013*, 339-340, Lancaster University, Lancaster, United Kingdom), July.

Herrero-Zorita, C., Campillos-Llanos, L. and Moreno-Sandoval, A. (2014). Collecting and POS-tagging a lexical resource of Japanese biomedical terms from a corpus. *Procesamiento del Lenguaje Natural*, 52, In press.

Hollmann, W. (2006). Passivisability of English periphrastic causatives. In *Corpora in Cognitive Linguistics. Corpus-based approaches to syntax and lexis* (pp. 101–125). Berlin/Heidelberg/New York: Mouton de Gruyter.

Howell, L. J. and Morimoto, H. (2004). Kanji Networks. Free Online Kanji Etymology Dictionary. An Etymological Dictionary of Chinese Character Interpretations. http://www.etymonline.com/abbr.php. Accessed 15/03/2013.

Ibarretxe-Antuñano, I. (2006). Estudio lexicológico de las onamatopeyas vascas: el euskal onomatopeien hiztegia: euskara-ingelesera-gaztelania. *Fontes Linguae Vasconum: Studia et Documenta*, *38*(101), 147–162.

Ibarretxe-Antuñano, I. (2010). Lexicografía y Lingüística Cognitiva. *Revista Española de Lingüística Aplicada*, (23), 195–214.

Irwin, M. (2011). *Loanwords in Japanese*. Amsterdam: John Benjamins Publishing.

Izumi, Y., and Isozumi, K. (2001). Modern Japanese medical history and the European influence. *The Keio journal of medicine*, *50* (2), 91-99.

Jackendoff, R. (2002): *Foundations of Language*. Oxford University Press, Oxford, New York

Jiang, H. (2012). A corpus-based approach to the study of collocation and cognitive perspective to semantic interpretation on adjectives in Japanese language: A case study on "amai." *Foreign Language Teaching and Research*, *44*(6), 845–855.

Molina, C. (2008) Historical dictionary definitions revisited from a prototype theoretical standpoint. Annual Review of Cognitive Linguistics (6).

Moreno-Cabrera, J. C. (1997) Introducción a la Lingüística. Enfoque Tipológico y Universal. Madrid: Síntesis.

Moreno-Sandoval, A., Campillos-Llanos. (2013). Design and Annotation of MultiMedica – A Multilingual Text Corpus of the Biomedical Domain. Procedia - Social and Behavioral Sciences, 95 (25): 33-39.

Soler, H. H. (2008). A metaphor corpus in business press headlines. *Ibérica: Revista de La Asociación Europea de Lenguas Para Fines Específicos ( AELFE )*, (15), 51–70.

The British National Corpus, version 3 (BNC XML Edition). (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Türker, E. (2013). A corpus-based approach to emotion metaphors in Korean: A case study of anger, happiness, and sadness. *Review of Cognitive Linguistics*, *11*(1), 73–144. doi:10.1075/rcl.11.1.03tur

Turner, M. (2007) Conceptual integration, in Geeraerts, D. and Cuyckens, H. (eds.) *The Oxford Handbook of Cognitive Linguistics.* Oxford: Oxford University Press.

Ungerer, F. (2007) Word-formation, in Geeraerts, D. and Cuyckens, H. (eds.) *The Oxford Handbook of Cognitive Linguistics.* Oxford: Oxford University Press.

Wulff, S. (2006). Go–V vs. go–and–V in English: a case of constructional synonymy? In *Corpora in Cognitive Linguistics. Corpus-based approaches to syntax and lexis* (pp. 101–125). Berlin/Heidelberg/New York: Mouton de Gruyter.

## 9. Appendices.

A. Frequencies and definitions of English terms in corpus containing *-gram*, *-graph* and *-graphy*.

| *-gram* | Raw Frequency | Normalised Frequency | Definition |
|---|---|---|---|
| Histogram | 55 | 14.39 | A graph in which values found in a statistical study are represented by vertical bars or rectangles |
| Autoradiogram | 14 | 3.66 | Another term for autoradiograph |
| Electrocardiogram | 13 | 3.40 | A graphic tracing of the variations in electrical potential caused by the excitation of the heart muscle and detected at the body surface. |
| Electroencephalogram | 8 | 2.09 | A recording of the potentials on the skull generated by currents emanating spontaneously from nerve cells in the brain, with fluctuations in potential seen as waves. |
| Interferogram | 7 | 1.83 | The representation in a diagram of the pattern formed by a wave interference |
| Angiogram | 5 | 1.31 | A radiographic image of a blood vessel after injection of a radiopaque contrast medium |
| Idiogram | 4 | 1.05 | A drawing or photograph of the chromosomes of a particular cell |
| Chromatogram | 3 | 0.78 | The record produced by the separation of gaseous substances or dissolved chemical substances moving through a column of absorbent material that filters out the various absorbates in different layers. Also: the record produced by chromatography. |
| Mammogram | 3 | 0.78 | A radiograph of the breast. Also: an x-ray image of the breast produced by mammography. |
| Tympanogram | 3 | 0.78 | A graphic representation of the relative compliance and impedance of the tympanic membrane and ossicles of the middle ear obtained by tympanometry |
| Arteriogram | 2 | 0.52 | An x-ray film of an artery injected with a radiopaque contrast medium |
| Electropherogram | 2 | 0.52 | A recording of the separated components of a mixture produced by electrophoresis |
| Electroretinogram | 2 | 0.52 | A graphic record of the electrical activity of the |

| | | | |
|---|---|---|---|
| | | | retina. |
| Sonogram | 2 | 0.52 | An image, as of an unborn foetus, produced by ultrasonography. Also called *echogram*, *ultrasonogram* |
| Engram | 1 | 0.26 | In the mnemic hypothesis, a physical change or memory trace made on the central nervous system of an organism as a result of experience or the repetition of stimuli |
| Spirogram | 1 | 0.26 | A tracing or graph of respiratory movements. Also, the tracing made by the spirograph. |

| ***-graph*** | **Raw Frequency** | **Normalised Frequency** | **Definition** |
|---|---|---|---|
| Micrograph | 33 | 8.63 | An instrument used to record very minute movements by making a greatly magnified photograph of the minute motions of a diaphragm |
| Radiograph | 18 | 4.71 | An image produced on a radiosensitive surface by radiation other than visible light, as by x-rays passed through an object |
| Photomicrograph | 13 | 3.40 | A photograph of an object as seen through an ordinary light microscope |
| Autoradiograph | 13 | 3.40 | An image recorded on a photographic film or plate produced by the radiation emitted from a specimen, such as a section of tissue, that has been treated or injected with a radioactively labeled isotope. |
| Chromatograph | 10 | 2.62 | The apparatus used in chromatography |
| Spectrograph | 5 | 1.31 | An instrument used to display or record spectra, as from electromagnetic or sound waves |
| Barograph | 1 | 0.26 | An instrument that continually monitors barometric pressure and records pressure changes on paper |
| Electro-cardiograph | 1 | 0.26 | An instrument used in the detection and diagnosis of heart abnormalities that measures electrical potentials on the body surface and generates a record of the electrical currents associated with heart muscle activity. |
| Magnetograph | 1 | 0.26 | A device for detecting and recording variations in the intensity and direction of magnetic fields |
| Oscillograph | 1 | 0.26 | A device that records oscillations, as of an electric current and voltage |

| | | | |
|---|---|---|---|
| Radiomicrograph | 1 | 0.26 | An enlarged x-ray photograph, used to study small details |
| Spirograph | 1 | 0.26 | An instrument for registering the depth and rapidity of respiratory movements |

| ***-graphy*** | **Raw Frequency** | **Normalised Frequency** | **Definition** |
|---|---|---|---|
| Chromatography | 112 | 29.30 | A method of separating and identifying the components of a complex mixture by differential movement through a two-phase system, in which the movement is effected by a flow of a liquid or a gas (mobile phase) which percolates through an adsorbent (stationary phase) or a second liquid phase |
| Stratigraphy | 57 | 14.91 | See tomography |
| Autoradiography | 46 | 12.04 | A photographic recording of radiation from radioactive material, obtained by placing the surface of the radioactive material in close proximity to a photographic emulsion |
| Topography | 42 | 10.99 | The description of the regions of the body or of a body part, especially the regions of a definite and limited area of the surface |
| Ultrasonography | 32 | 8.37 | The process of imaging deep structures of the body by measuring and recording the reflection of pulsed or continuous high-frequency sound waves. Valuable in many medical situations, including the diagnosis of fetal abnormalities, gallstones, heart defects, and tumors. |
| Radiography | 19 | 4.97 | The production of shadow images on photographic emulsion through the action of ionizing radiation. The image is the result of the differential attenuation of the radiation in its passage through the object being radiographed |
| Tomography | 23 | 6.02 | The recording of internal body images at a predetermined plane by means of a tomograph |
| Cardiotocography | 16 | 4.19 | Method of monitoring and recording foetal heart rate and uterine contractions during pregnancy and labour, allowing for assessment of foetal response and well-being |
| Angiography | 13 | 3.40 | The radiographic visualization of the internal anatomy of the heart and blood vessels after the intravascular introduction of radiopaque contrast medium |

| | | | |
|---|---|---|---|
| Echocardiography | 9 | 2.35 | Recording of the position and motion of the heart walls or internal structures of the heart by the echo obtained from beams of ultrasonic waves directed through the chest wall |
| Arteriography | 8 | 2.09 | A method of radiologic visualization of arteries performed after a radiopaque contrast medium is introduced into the bloodstream or into a specific vessel by injection or through a catheter |
| Electroencephalography | 8 | 2.09 | The recording of changes in electric potential in various areas of the brain by means of electrodes placed on the scalp or on or in the brain itself |
| Photomicrography | 7 | 1.83 | The process of taking photographs through a microscope |
| Electrocardiography | 5 | 1.31 | The making of graphic records of the variations in electrical potential caused by electrical activity of the heart muscle and detected at the body surface, as a method for studying the action of the heart muscle |
| Mammography | 5 | 1.31 | Radiographic examination of the breasts for diagnostic purposes |
| Urography | 4 | 1.05 | The radiographic examination of the urinary system. A radiopaque substance is injected, and radiographs are taken as the substance is passed through or excreted from the part of the system being studied. |
| Fluorography | 3 | 0.78 | The photographic recording of fluoroscopic images on small films, using a fast lens; used in mass radiography of the chest |
| Phlebography | 2 | 0.52 | The radiographic examination of veins injected with a radiopaque contrast medium. |
| Ventriculography | 2 | 0.52 | The radiographic examination of a ventricle of the heart after injection of a radiopaque contrast medium |
| Electromyography | 1 | 0.26 | The recording and study of the electrical properties of skeletal muscle |
| Immunoscintigraphy | 1 | 0.26 | Scintigraphic imaging of a lesion using radiolabeled monoclonal antibodies or antibody fragments specific for antigen associated with the lesion |
| Plethysmography | 1 | 0.26 | The measurement of changes in the volume of organs or other body parts, particularly those changes resulting from blood flow |
| Pyelography | 1 | 0.26 | X-ray photography of the pelvis of the kidney and associated structures after injection with a |

| | |
|---|---|
| | radiopaque dye. |

B. Frequencies and definitions of Japanese terms in corpus containing the counterparts for *-gram*, *-graph* and *-graphy*.

| **Morpheme equal to *-gram*** | **Medical Term** | **Raw Frequency** | **Normalised Frequency** | **Translation** |
|---|---|---|---|---|
| グラム | オートラジオグラム | 1 | 1.00 | Autoradiogram |
| | クロマトグラム | 2 | 2.00 | Chromatogram |
| | ガリウムシンチグラム | 1 | 1.00 | Gallium scintigram |
| | シンチグラム | 10 | 10.02 | Scintigram |
| | オージオグラム | 2 | 2.00 | Audiogram |
| | ヒストグラム | 4 | 4.01 | Histogram |
| 図 | 筋電図 | 3 | 3.01 | Electromyogram |
| | 蝸電図 | 2 | 2.00 | Electrocochleogram |
| | 聴力図 | 1 | 1.00 | Audiogram |
| 写真 | X線写真 | 16 | 16.04 | Roentgenogram |
| | 胸部X線写真 | 6 | 6.01 | Chest roentgenogram |
| 像 | エコー像 | 8 | 8.02 | Echogram |
| | 超音波像 | 8 | 8.02 | Ultrasonogram |
| | 断層像 | 5 | 5.01 | Tomogram |
| | X線像 | 10 | 10.02 | Roentgenogram |

| **Morpheme equal to *-graph*** | **Medical Term** | **Raw Frequency** | **Normalised Frequency** | **Translation** |
|---|---|---|---|---|
| グラフ | マイクロライノグラフ | 1 | 1.00 | Microrhinograph |
| | レーダーグラフ | 2 | 2.00 | Radargraph |

| | | | | |
|---|---|---|---|---|
| 写真 | X線写真 | 21 | 21.05 | Radiograph (or roentenogram) |
| | 胸部X線写真 | 6 | 6.01 | Chest radiograph (or roentenogram) |

| **Morpheme equal to *-graphy*** | **Medical Term** | **Raw Frequency** | **Normalised Frequency** | **Translation** |
|---|---|---|---|---|
| グラフィー | シンチグラフィー | 18 | 18.04 | Scintigraphy |
| | イムノクロマトグラフィー | 13 | 13.03 | Immunochromatography |
| | クロマトグラフィー | 11 | 11.02 | Chromatography |
| | カラムクロマトグラフィー | 2 | 2.00 | Column chromatography |
| | ガスクロマトグラフィー | 4 | 4.01 | Gas chromatography |
| | オートラジオグラフィー | 1 | 1.00 | Autoradiography |
| | カラードプラアンギオグラフィー | 1 | 1.00 | Color Doppler angiography |
| | サーモグラフィー | 3 | 3.01 | Thermography |
| 造影 | 腹部血管造影 | 1 | 1.00 | Abdominal angiography |
| | 動脈造影 | 32 | 32.07 | Arteriography |
| | 門脈造影 | 28 | 28.06 | Portography |
| | 血管造影 | 1 | 1.00 | Angiography |
| 検査 | 超音波検査 | 86 | 86.19 | Ultrasonography |
| | 腹部超音波検査 | 92 | 92.20 | Abdominal ultrasonography |
| | X線検査 | 16 | 16.04 | Radiography |
| | 睡眠ポリグラフ検査 | 8 | 8.02 | Polysomnography |
| | エコー検査 | 4 | 4.01 | Echography |
| | 心エコー検査 | 3 | 3.01 | Echocardiography |
| | 重心動揺検査 | 5 | 5.01 | Posturography |

| | | | | |
|---|---|---|---|---|
| 図検査 | 筋電図検査 | 1 | 1.00 | Electromyography |
| | 蝸電図検査 | 5 | 5.01 | Electrocochleography |
| 造影検査 | 血管造影検査 | 15 | 15.03 | Angiography |
| 撮影 | 脳血管撮影 | 1 | 1.00 | Cerebral angiography |
| | 血管撮影 | 2 | 2.00 | Angiography |