New Perspectives in Science Education
International Conference
NEW PERSPECTIVES in SCIENCE EDUCATION


Popularization Strategies in Public Health Discourse: a Corpus-based Analysis in Portuguese and Chinese

Huanqi Sun[1], Katrin Herget[2], João Paulo Silvestre[3]

University of Aveiro, PhD candidate, Centre for Languages, Literatures and Cultures, Portugal[1]
University of Aveiro, Centre for Languages, Literatures and Cultures, Portugal[2, 3]

Abstract

An increasing number of laypeople are turning to the vast information available on the Internet to self-educate in public health matters. For instance, they seek advice from experts in specific subject areas to understand how Covid-19 spreads or to clarify uncertainties about physiological reactions during pregnancy. Additionally, they actively engage in information exchange through magazines, websites, journalists, and other platforms on relevant topics. All of these efforts reflect laypeople's attempts to break down barriers in expert-layperson communication. In this context, the cultivation of scientific literacy emerges as a crucial skill, enabling individuals to critically assess scientific information and make informed decisions.

The vertical structure of medical discourse in Language for Specific Purposes (LSP) primarily distinguishes between internal communication and external discourse. External discourse, which is patient-oriented, involves doctor-patient interactions and communication with a broader audience [1]. The objective of discourse popularization is to ensure that scientific information becomes accessible to a wider audience. Therefore, it is imperative to adjust the language used based on the level of popularization required [2].

This paper explores strategies for popularizing scientific discourse in the field of public health in both Portuguese and Chinese, highlighting effective ways to enhance scientific literacy in these language communities. The corpus analysis is carried out using the corpus manager and text analysis tool, Sketch Engine [3].

Keywords: *Public health, discourse, popularization, comparable corpora, strategies*

1. Introduction

In tandem with societal advancement, an increasing diversity of food choices has become widely available. The promotion of guidelines for healthy eating is of practical importance to the public. For example, informative advice on food selection, strategies for preventing diseases through sensible dietary choices, and considerations for dietary restrictions during illness, among others. The complexity of food choices and fast-paced modern life underline the importance of clear guidance on healthy eating. This information empowers people to make informed choices, preventing long-term health problems. This paper aims to examine discourse strategies in science popularization texts, with a specific focus on analysing the utilization of discourse markers. Discourse markers (DMs) are pragmatically defined as linguistic elements that serve a connective function in constructing texts [4].

The study is designed to analyze paraphrastic reformulation markers in popular science texts in both Portuguese and Chinese. To accomplish this, we will investigate reformulation as a strategy employed in the process of science popularization.

The definition of reformulation aims to repair formulation problems and guarantee intercomprehension, focusing on cases of self-reformulation: “Reformulation is a meta-discursive operation whereby the speaker reworks an utterance (or part of it) with the intention of making their discourse more intelligible, reducing any risk of misunderstanding on the part of the interlocutor” [5] (p. 2694) [own translation].

Reformulation is one of the strategies of science popularization, which is considered a form of simplification [1]. In an attempt to systematize reformulation markers, the authors consider reformulation as an interpretative process, serving to paraphrase, complete, or correct [6] (pp. 320-321). Two types of reformulation can be distinguished: paraphrastic (relationship of equivalence between the source utterance and the reformulated utterance) and non-paraphrastic (correction or rectification) [7]. In this paper, we analyse paraphrastic reformulation markers.

2. Reformulation markers

Research on discourse markers (DM) emerged in the mid-20th century. DM have the function of organising the textual structure, signalling either a return to what has been said in order to clarify or summarise it, or a planning strategy. Based on previous studies([4], [5] [6], [8] and [9]), we propose an analysis of paraphrastic markers of reformulation (marking of an equivalence relation). When the speaker rephrases a fragment of his or her speech, he or she introduces a clarification that makes what was said earlier more precise. The most common markers for signalling this type of reformulation include expressions that are summarized in the table below:

Reformulation markers in Chinese	Reformulation markers in Portuguese
即 Pinyin (PY): (jí) Literal translation (LT): i.e. PT: nomeadamente ENG: namely	-
换言之 (huàn yán zhī) LT: change words it PT: por outras palavras ENG: in other words	-
换句话说/换句话讲 (huàn jù huà shuō / huàn jù huà jiǎng) LT: change sentence words explain / change sentence words explain ENG: in other words	ou seja noutros termos, dito de outro modo, por outras palavras
也就是说/这就是说 (yě jiù shì shuō / zhè jiù shì shuō) LT: also that's it yes explain / this that's it yes explain ENG: that is to say	isto é

这(就)意味着 (zhè (jiù) yì wèi zhe) LT: this (that's) mean taste with ENG: meaning / that is	quer dizer

Table 1. Most common paraphrastic markers of reformulation

In the Chinese language, reformulation markers *换句话说/换句话讲* (huàn jù huà shuō / huàn jù huà jiǎng) (in. *in other words*) corresponds to pt. *ou seja*, *noutros termos*, *dito de outro modo*, *por outras palavras*.

3. Method

For the present study, we compiled two corpora, each consisting of thirty popular science texts in both Portuguese and Chinese, specifically in the field of Health Sciences with a focus on well-being and diets. Both corpora will be analyzed using the Sketch Engine tool. The investigation of the selected articles in both languages will involve: firstly, identifying paraphrastic reformulation markers; secondly, determining the overall frequency of reformulation markers; and thirdly, analyzing the function of these reformulation markers based on context.

We compiled two comparable corpora, with a focus on ensuring comparability based on criteria such as chosen topics and text length. This preliminary analysis aims to compare Chinese and Portuguese texts, identifying similarities and differences in the use of reformulation markers. We have integrated both qualitative and quantitative analyses to provide a more comprehensive approach to data examination. To identify the texts in the analysis, numbers are used to present and distinguish the articles. For example, Portuguese texts are labelled as PT1, PT2, PT3, and Chinese texts are labelled as CH1, CH2, CH3, and so forth. Since in Portuguese there is no equivalent for the Chinese reformulation markers *即(jí) (namely, i.e.)* and *换言之 (huàn yán zhī) (to put it differently)*, as shown in Table 1, we focused our analysis on the following markers: *ou seja*, *noutros termos*, *dito de outro modo*, *por outras palavras*, *isto é*, *quer dizer* and its corresponding Chinese equivalents.

As far as corpus compilation is concerned, we gathered Portuguese popular science texts from free online magazines, such as *Revista Saúde* and *Revista Prevenir*. However, our search for accessible Chinese popular science texts on Google proved challenging due to limited open-access resources. Consequently, we adopted an alternative approach to search for texts. We obtained permission from the "文献云图书馆" (Literature Cloud Library), a website containing numerous Chinese academic papers, as well as texts from magazines and journals. Additionally, we searched for texts in Chinese libraries, such as "大众医学" (Popular Medicine). Furthermore, segments of texts containing reformulation markers in both Portuguese and Chinese from these sources were extracted and compiled into a corpus.

The popular science texts in the Portuguese language were collected from the following sources: the website of the Portuguese pharmacies, from which we obtained access to *Revista Saúda* [10] (12 texts) in the health field, covering topics such as Cancer Prevention and Considerations for Diabetics, as well as subfields, like well-being and diets. Additionally, texts were sourced from *Revista Prevenir* (6 texts) [11]; *Vida Ativa* (1 text) [12]; *Saúda e Bem Estar* (2 texts) [13]; *Visão Saúde* (1 text) [14]; *Viver Saudável* (2 texts) [15]; *Agroportal* (1 text) [16]; *Saúde Actual* (5 texts, available in print) [17]. The magazines are generally available in open access, except for *Revista Actual*. In the Chinese language popular science texts were sourced from the website of the "文献云图书馆" (Literature Cloud Library) (36 texts) [18] and "大众医学" (Popular Medicine (4 texts, available in print) [19]. The collected texts in both languages cover the period from 2016 to 2023.

4. Results and Discussion

In the following, we present and discuss results obtained from both the Chinese and Portuguese corpora.

5. Analysis of reformulation markers

Upon analyzing the frequencies of reformulation markers using the Concordance function of the Sketch Engine tool, the following data were obtained:

Corpus PT	Corpus CH	Frequency			
		ou seja	noutros termos, ditos de outro modo, por outras palavras	isto é	quer dizer
		换句话说 (huàn jù huà shuō) LT: change sentence words explain / change sentence words explain		也就是说 (yě jiù shì shuō) LT: also that's it yes explain / this that's it yes explain	这(就)意味 (zhè (jiù) yì wèi zhe)
PT20		4	0	0	0
PT1, PT3		3	0	0	0
PT2, PT15, PT19, PT23, PT26, PT29, PT30		2	0	0	0
PT4, PT5, PT6, PT7, PT8, PT9, PT10, PT11, PT12, PT14, PT16, PT17, PT18, PT21, PT22, PT24, PT25, PT27, PT28	CH8，CH9，CH10，CH1，CH12，CH13，CH15，CH17，CH18, CH19, CH22, CH23, CH24, CH25, CH27, CH28	1	0	0	0
PT13		1	0	1	0
PT18	CH1, CH2, CH3, CH4, CH5, CH6, CH7, CH14, CH16, CH21, CH26, CH30	0		1	0
	CH29	0		3	
	CH20	0		2	

Table 2. Frequency of reformulation markers in both corpora

According to the Table 2, the analyzed corpora did not contain reformulation markers 这(就)意味着 (zhè (jiù) yì wèi zhe, quer dizer). In the Portuguese corpus, no instances of noutros termos, dito de outro modo, por outras palavras were found. The reformulation marker ou seja revealed 44

occurrences, significantly surpassing the occurrences of isto é, which totalled only 1.In the Chinese corpus, the reformulation markers
也就是说 (yě jiù shì shuō) (isto é) and 换句话说 (huàn jù huà shuō) (ou seja) appeared almost equally, with 17 and 16 occurrences, respectively. There were no instances of 换句话讲 (huàn jù huà jiǎng) or 这就是说 (zhè jiù shì shuō) in the Chinese corpus.

5.1 Analysis of reformulation markers: *ou seja* (*huàn jù huà shuō*) and *isto é* (*yě jiù shì shuō*)

The discourse markers *isto é* and *ou seja* share a similar meaning, translating into English as *that is* or *in other words*. However, they are employed in slightly distinct contexts in the Portuguese language. Considering the length of the paper, we selectively analyze text fragments in both languages, with reformulation markers in text highlighted in bold.

PT13: (...) É nos vasos por onde circula o sangue que se vão dar os principais efeitos das substâncias tóxicas que o tabaco contém, danificando o endotélio vascular, *isto é*, a fina camada celular que reveste o interior dos vasosÉ suspender totalmente o tabaco e moderar o consumo de álcool. Se isso for feito antes de serem estabelecidas lesões irreversíveis, pode haver melhoria das queixas. *Ou seja*, quanto mais cedo se deixar o vício do tabaco e o abuso do álcool, melhores são as perspectivas de recuperação, não só sexuais, como do sistema cardiovascular e de outros sistemas. In segment PT13, *isto é* is employed to elucidate the medical term *endotélio vascular*, ensuring enhanced reader comprehension by providing a simplified explanation. And *ou seja* is employed to reiterate the positive health effects of quitting smoking and limiting alcohol consumption. It emphasizes various facets of this information, offering a different perspective on a prior statement to enhance comprehension. This suggests a deliberate rephrasing or reorganization of words to better convey the author's intention. Overall, both expressions are used in the text to ensure that the reader's understanding of the previous statement is more accurate and comprehensive.

PT20: (...) O colesterol faz parte da membrana celular, se existir falta de colesterol, quando renova as suas células, estas podem apresentar rigidez na membrana e como tal as trocas celulares ficarão comprometidas. Pesquisas recentes indicam que o colesterol pode atuar como antioxidante, *ou seja*, protege as células contra efeitos nocivos dos radicais livres. (...) Quanto mais adipócitos encher mais vai aumentar de volume e peso, *ou seja*, engordar! Isto é o que acontece quando você faz exercício e os seus músculos necessitam de energia. É por isso que o exercício intenso emagrece! Durante o exercício acontece um processo chamado glicogénese, *ou seja*, a glicose é transformada em glicogénio para ser utilizado como combustível. Se o glicogénio começa a baixar, o organismo inicia um outro processo, a neoglicogénese, *ou seja*, vai às reservas buscar estes triglicéridos e utiliza as moléculas de gordura como energia e o glicerol como glicose, deste modo volta a haver energia e você fica mais magro.
The first *ou seja* is used as a rephrasing and explanation of the preceding text to emphasise the antioxidant role of cholesterol. It further clarifies the role of cholesterol and makes it easier for the reader to understand its role in cell protection. Additionally, *ou seja* introduces new insights from recent research about the antioxidant properties of cholesterol, ensuring a more comprehensive understanding for the reader. The second *ou seja* is used as a rephrasing of the previous text to emphasise the relationship between fat cell filling and weight gain. It provides a more detailed and

visual explanation to ensure that the reader has a clearer understanding of the mechanisms of weight gain. The third *ou seja* is used to explain the *glicogénese* process mentioned earlier, in order to summarise and emphasise the nature of the process. It ensures that the reader understands the mechanism by which glucose is converted into liver glycogen during exercise. The fourth *ou seja* is used to explain the process of *neoglicogénese* mentioned earlier to provide a more detailed description. It emphasises how the body uses fat reserves to produce energy when glycogen is low, and promotes a fuller understanding of the physiological process of weight loss. *Isto é* clarifies the previous statement, detailing the process of filling adipocytes and its relationship with weight gain.

CH23: (...) It can be seen that the culprit of hyperlipidaemia is improper diet, and the key to preventing and controlling hyperlipidaemia is also diet, *换句话说 (huàn jù huà shuō, in other words)*, improper diet will lead to hyperlipidaemia [own translation]. [The function of *换句话说* is reformulation and simplification of the last sentence]
In this context, the use of *huàn jù huà shuō (in other words)* is to summarise and emphasise the preceding text in a more concise and straightforward way, which is also a form of reformulation. The author uses *换句话说 (huàn jù huà shuō)* to emphasise that the main cause of hyperlipidaemia is improper diet and puts the key to preventing hyperlipidaemia on diet. This style of presentation helps to highlight the direct effect of diet on hyperlipidaemia and conveys complex information to the reader in a clearer way.

CH14: (...) However, during the production of infant formula, the manufacturer will make regular "breastfeeding adjustments" to goat's milk or cow's milk, which includes adjusting the ratio of whey and casein proteins in the raw materials so that it reaches the ratio of proteins in breastmilk. *也就是说 (yě jiù shì shuō)*, both cow's milk and goat's milk have protein ratios that are closer to breast milk, but are ultimately adjusted to be close to breast milk, and the difference is very small, so there is no such thing as goat's milk formula being closer to breast milk [own translation]. [explanation of *breastfeeding adjustments*]
Through the phrase *也就是说 (yě jiù shì shuō)*, the author emphasises that during the production of infant formula, the protein ratio of both cow's milk and goat's milk will be adjusted to be close to that of breast milk, so that readers will understand that the difference between goat's milk powder and cow's milk powder in terms of protein composition is very small. This helps to correct possible misconceptions and gives the reader a fuller understanding of the process of preparing speciality milk and infant formula.

5.2 Some other results

Many Chinese reformulation markers, such as *换句话说 (huàn jù huà shuō, ou seja)*, are prominently present in the first paragraphs of popular science texts. Specifically, these markers are notably featured in texts CH11, CH12, CH13, CH22, CH23, CH24, CH27, and CH28. Similarly, the reformulation markers *也就是说 (yě jiù shì shuō, isto é)* are prevalent in the opening paragraphs of various popular science texts, including those in CH1, CH2, CH3, CH20, CH21, CH26, and CH30. However, in the Portuguese corpus, instances of the reformulation marker *ou seja* only appear in the first paragraph in texts PT2, PT6 and PT30. Perhaps this phenomenon is something worth exploring in future studies.

6. Conclusions

The discourse marker *ou seja* not only serves the function of reformulation: on the one hand, *ou seja* is usually employed to rephrase or express something in a different way to enhance understanding and be used for discourse strategies that imply a reformulation of the speech. Nevertheless, it can also be used to introduce new information for clarification purposes, as the pragmatic distinction is not always easy to establish. The discourse marker *isto é* is used to offer additional clarification or explanation of a preceding statement. This technique enhances logical coherence and overall text readability. In summary, the utilization of *isto é / yě jiù shì shuō* in these contexts aims to improve clarity, ensuring that the reader precisely grasps the author's intended message. Moreover, studying reformulation markers contributes to a better understanding of popular science texts and enhances scientific literacy. This study, albeit small, provides insight into paraphrastic reformulation markers in Portuguese and Chinese texts in the field of Health Sciences. For more details and results regarding the functions of *ou seja* and *isto é*, further analysis of additional texts may be necessary in the future.

References

[1] Brand, C. (2008). Lexical processes in scientific discourse popularisation: a corpus-linguistic study of the SARS coverage. Peter Lang, pp.27-28.

[2] López Orellana, M. (2012). Popularising scientific discourse. Quaderns de Filologia. Estudis lingüístics. Vol. XVII (2012), pp. 83-96.

[3] Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. Lexicography, 1, pp. 7-36.

[4] Hong, W., & Yunfeng, G. (2004). 重构性话语标记语的英汉对比分析(A Comparative English-Chinese Analysis of Reconstructive Discourse Markers). *Journal of Zuoling College*, (012), pp. 160-162.

[5] Lopes, A. C. M. & Carrilho, E. (2020). Discurso e Marcadores Discursivos. In E. B. Paiva Raposo et al. (orgs.) *Gramática do Português.* (pp. 2667-2730). Vol. III. Lisboa: Fundação Calouste Gulbenkian.[6] Gülich E. & Kotschi. T (1983). Les marqueurs de la reformulation paraphrastique. In J. Moeschler Ed). *Connecteurs pragmatiques et structure du discours. Cahiers de linguistique française*, Univ. de Genève, pp. 305-351.

[6] Gülich, E., & Kotschi, T. (1983). Les marqueurs de la reformulation paraphrastique. In Connecteurs pragmatiques et structure du discours; actes du 2ème Colloque de Pragmatique de Genève (7-9 mars 1983) Vol. 5, pp. 320-321.

[7] Rossari, C. (1984). Les opérations de reformulation. Peter Lang.

[8] Shulin, Y. (2012). 现代汉语话语标记研究(A Study of Discourse Markers in Modern Chinese). China Social Science Press. ISBN: 7516118109, 9787516118108

[9] Xiao, R. (2010). Idioms, word clusters, and reformulation markers in translational Chinese: Can "translation universals" survive in Mandarin? In *The International Symposium on Using Corpora in Contrastive and Translation Studies. Ormskirk: Edge Hill University*, pp.1-40.

[10] https://www.revistasauda.pt/Revistas/Revista_Sauda/Pages/default.aspx

[11] https://www.prevenir.pt/

[12] https://www.vidaativa.pt/

[13] https://saudebemestar.com.pt/

[14] https://visao.pt/visaosaude/

[15] https://www.viversaudavel.pt/

[16] https://www.agroportal.pt/
[17] https://saudeactual.pt/wp/
[18] http://wen.xstsg.top
[19] https://www.popumed.com/