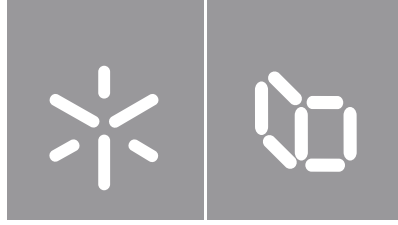


Universidade do Minho
Escola de Letras, Artes e Ciências Humanas

Yulianti Suryawardhani Yudo

**Analyzing Portuguese Loanwords
within a Dictionary of Indonesian Language**

Europäischer Master für Lexikographie
EMLex
European Master in Lexicography



Universidade do Minho

Escola de Letras, Artes e Ciências Humanas

Yulianti Suryawardhani Yudo

**Analyzing Portuguese Loanwords
within a Dictionary of Indonesian Language**

Master Thesis
European Master in Lexicography

Supervised by
Prof. Dr. Álvaro Iriarte Sanromán
Dr. Dominik Brückner

July 2022

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



**Atribuição-NãoComercial-Compartilhalgual
CC BY-NC-SA**

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Acknowledgements

First and foremost, I would like to thank EMLex and the entire consortium for giving me the opportunity to continue with my studies in Europe with this unique program.

Secondly, I would like to thank both of my supervisors, Prof. Dr. Álvaro Iriarte Sanromán from Universidade do Minho in Braga, Portugal, and Dr. Dominik Brückner from Institut der Deutschen Sprache in Mannheim, Germany, who have pointed me to the right direction when I was confused and lost, and also shared with me so much of their knowledge, academically and otherwise.

I would also like to thank Prof. Dr. Idalete Maria da Silva Dias, the coordinator of EMLex in Universidade do Minho and our mother while in Braga. Her dedication, love and kindness really made me feel at home. She always came to our rescue when things got tough and I can't appreciate her enough for that.

Of course, I can't forget the brother and sister from different parents that I've gained during this program, David Huerta Meza and Snizhana Umanets. They have no idea how thankful I am to have met them. We shared this adventure and went through various ups-and-downs in the presence of one another. Hopefully this bond will last beyond the time we've shared together these past few years.

To my real family: My mother Hertiningrum who never ceased to remind me to take care while being away from home and always prayed for my safety, as well as updated me with every little thing at home. My father Hindaryudo Soekarman, who's always supported my choices and encouraged me to go beyond my comfort zone, as well as shared words of wisdom and discussed about anything under the sun with me. Also my brothers, Sakti Hananto Yudo and Kuku Hario Yudo; we might not have communicated much during my time away from home, but they are always in my mind and heart, and so are their little families.

During my stay in Portugal, I've been involved in the Association of Indonesian Students in Portugal (PPI Portugal). The members welcomed me with open arms when I first contacted them before my arrival, helped out with any question or trouble I've had, encouraged me to become part of the association body, then secretary, and finally appointed me as president since October 2021. It is truly an honor to have become an integral part of this close-knit family as it has widened my horizon and upgraded me in the process.

In several parts of this thesis, I mentioned the work of Daya Negri Wijaya. He isn't a mere scholar whose relevant work I cited, but he's also been a great friend and respected senior while he's doing his PhD in Universidade do Porto. I appreciate the insights, the discussions and the friendship he's shared with me.

To my flatmates Alice Mevis and Emilio Jerez, who have not only kept me entertained and full, but also talked about many different aspects of life with and given me insights from their points of view. Obrigada!

Last but not least, I'd like to thank myself. Me, Yulianti Suryawardhani Yudo, for finishing this program and thesis in one piece; for the perseverance, the patience, the sweat and tears that were poured during the process. It was a very life-changing period, but I've made it! Here's to more adventures in life!

Analyzing Portuguese Loanwords within a Dictionary of Indonesian Language

Abstract

This thesis studies the Portuguese loanwords in Indonesian language. Portuguese is one of the languages that has influenced and enriched Indonesian since the 16th century, and now some loanwords have become part of the vocabulary. However, little has been studied in comparison to other linguistic influences and therefore, through this study, we would like to analyze those loanwords from Portuguese as preliminary study, before ultimately be able to create a dictionary of loanwords with Portuguese, since in Indonesia, we do not have one that is comprehensive.

The analysis is done through compiling the already available glossaries of Portuguese loanwords. The resulted compilation is then analyzed for its frequency and semantic field. This process is done with the help of a corpus and some available online resources related to loanwords, particularly in Indonesian, but also in general.

Finally, we try representing some select loanwords in a dictionary. For this, we additionally look up for some etymological information with the help of some historical Portuguese dictionaries, as well as Portuguese etymological dictionaries. This is a necessary addition, because after all, a loanwords dictionary would not be comprehensive enough without etymological information.

Keywords: Indonesian language, loanwords, Portuguese loanwords

Análise dos Empréstimos do Português num Dicionário de Indonésio

Resumo

Este trabalho trata dos empréstimos do português no indonésio. O português é uma das línguas que têm influenciado no indonésio, desde o seu contacto das duas línguas no século XVI. No entanto, esta influência não tem sido suficientemente estudada, se compararmos com outras línguas que têm influenciado o indonésio ao longo da história. Por essa razão, decidimos fazer este estudo, e também para, posteriormente, poder criar um dicionário de empréstimos do português no indonésio.

A análise dos empréstimos é feita usando glossários que já foram compilados em estudos prévios. A partir destes glossários foi compilada uma lista de empréstimos, que analisamos segundo a frequência e o campo semântico. Para isso, foram usados corpora e outros recursos ligados a empréstimos linguísticos, especialmente no indonésio, mas também em geral.

Finalmente, apresentamos exemplos ilustrativos de como os empréstimos do português no indonésio serão recolhidos no dicionário de empréstimos. Neste dicionário é imprescindível a recolha de informação etimológica, que, neste caso, é feita com a ajuda de um dicionário histórico do português e da informação etimológica recolhida num dicionário português, nos dicionários etimológicos do português e num dicionário etimológico hispânico.

Palavras-chave: empréstimo, empréstimos do português, indonésio

Table of Contents

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS	ii
Licença concedida aos utilizadores deste trabalho.....	ii
STATEMENT OF INTEGRITY	iii
Acknowledgements	iv
Abstract	v
Resumo	vi
Introduction	1
0.1. Objective of the Study	3
0.2. Research Question	3
0.3. Structure of the Study	4
Chapter 1: Theoretical Background	5
1.1. Language Borrowing	5
1.2. Loanwords in Indonesian Language	6
1.3. Portuguese Loanwords in Indonesian and Malay Language	8
Chapter 2: Methodology	11
2.1. Data Source	11
2.1.1. Loanword Online Resources	11
2.1.2. Glossaries	20
2.1.3. Historical and Etymological Dictionaries	20
2.2. Methodology	21
Chapter 3: Results & Discussion	23
3.1. Most Frequent Words in Indonesian Language	23
3.2. Frequency of Portuguese Loanwords	26
3.3. Connotations of Portuguese Loanwords	37
3.4. Semantic Fields of Portuguese Loanwords	40
3.4.1. Commerce	41
3.4.2. Entertainment	43
3.4.3. Warfare	45
3.4.4. Religion	46
3.4.5. Household	49
3.4.6. Clothing	50

Chapter 4: Dictionary Representation	54
4.1. Early Representation Idea	54
4.2. Representation Idea after Analysis	55
4.2.1. Palsu	57
4.2.2. Boneka	59
4.2.3. Armada	60
4.2.4. Sekolah	61
4.2.5. Meja	63
4.2.6. Saku	64
Conclusion	66
Bibliography	68
Appendix I	74

List of Figures

Figure 1 Result page for the entry Natal on Corpora Collection of University of Leipzig, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=Natal (retrieved on December 23, 2021)	12
Figure 2 Result page of the entry Natal on KBBI Daring, from https://kbbi.kemdikbud.go.id/entri/Natal (retrieved on December 23, 2021).....	13
Figure 3 Map of Indonesian as a recipient language and its donor languages on WOLD, from: https://wold.cld.org/language/27 (retrieved on January 17, 2022).....	15
Figure 4 List of loanwords in Indonesian on WOLD, in alphabetical order, from https://wold.cld.org/language/27 (retrieved on January 17, 2021).....	16
Figure 5 List of Portuguese loanwords in Indonesian on WOLD, from: https://wold.cld.org/language/27#recipient-map (retrieved on January 17, 2022).....	17
Figure 6 Information page of the entry meja on WOLD, from: https://wold.cld.org/word/71261839983027004 (retrieved on January 10, 2022)	18
Figure 7 Map of Indonesian as donor language based on WOLD, from: https://wold.cld.org/language/27 (retrieved on January 17, 2022).....	19
Figure 8 Chart of 50 most frequent words in Indonesian	25
Figure 9 https://kbbi.kemdikbud.go.id/Beranda/Statistik (retrieved on January 12, 2022)	32
Figure 10 https://kbbi.kemdikbud.go.id/Beranda/Pemutakhiran (retrieved on January 12, 2022).....	32
Figure 11 Chart of most frequent Portuguese loanwords	34
Figure 12 Rank of Portuguese loanwords	35
Figure 13 Word graph of the loanword kamar (room) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=kamar (retrieved on January 22, 2022)	38
Figure 14 Word graph of the loanword palsu (false) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=palsu (retrieved on January 22, 2022).....	39
Figure 15 Word graph of the loanword peluru (bullet) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=peluru (retrieved on January 22, 2022)	40
Figure 16 Semantic field of Portuguese loanwords in Indonesian: commerce	41
Figure 17 Word graph of the loanword tempo (time, speed) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=tempo (retrieved on January 21, 2022)	42

Figure 18 Semantic field of Portuguese loanwords in Indonesian: entertainment.....	43
Figure 19 Word graph of the loanword <i>legenda</i> (legend) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=legenda (retrieved on January 21, 2022)	44
Figure 20 Semantic field of Portuguese loanwords in Indonesian: warfare.....	45
Figure 21 Word graph of the loanword <i>armada</i> (fleet of ships) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=armada (retrieved on January 24, 2022)	46
Figure 22 Semantic field of Portuguese loanwords in Indonesian: religion	46
Figure 23 Semantic field of Portuguese loanwords in Indonesian: household	49
Figure 24 Word graph of the loanword <i>kampung</i> (village) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=kampung (retrieved January 25, 2022)	49
Figure 25 Semantic field of Portuguese loanwords in Indonesian: clothing	50
Figure 26 Semantic Fields of Portuguese Loanwords.....	52
Figure 27 Example of dictionary entry for a Portuguese loanword (Yudo, 2020)	54
Figure 28 Template of the microstructure of our dictionary of Portuguese loanwords.....	57

List of Tables

Table 1 The 50 most frequent Indonesian words	24
Table 2 Frequency of Portuguese Loanwords in Indonesian	29
Table 3 Portuguese loanwords among the 10,000 most frequent words based on KBBI Online entries	34
Table 4 Frequency of Indonesian names of the days of the week	36

Introduction

For centuries, the Indonesian archipelago was a hub for merchants coming from different places, meaning that they have different backgrounds, cultures and languages as well. These merchants, connected to a network of coastal commerce port-cities, spoke a common commerce language, which was Malay (Adam, 2020, p. 61). Nowadays, there are two standardized versions of Malay; Malaysian, the variety used in Malaysia, Singapore and Brunei, also called *Bahasa Malaysia*; and Indonesian, the variety used in Indonesia, also called *Bahasa Indonesia* (Haspelmath & Tadmor, 2009, p. 687).

The Portuguese were the first Europeans to explore the archipelago, starting the 16th century, controlling the trade of spices. According to Professor De Vries, “[t]he Portuguese introduced hitherto unknown everyday articles with their Portuguese names, such as foods, plants and clothes, and in addition they brought the Catholic faith with its own vocabulary.” (Jones, 2008, p. xxx¹)

In 17th century, the arrival of the Dutch pushed the Portuguese to the eastern part of the archipelago, although even there, the Dutch were able to put their influence, for example by not allowing the locals to keep their Catholic faith and forcing them to convert to Protestantism (Wijaya, Wahyudi, Umaroh, 2021, p. 81). However, during their occupancy in the archipelago, the Dutch also adopted Malay as a language of the colonial administration, making the language progressively standardized and disseminated among the educated native elites (Errington, 1986, p. 335).

It was the simplicity of this Malay² that made this language the choice of Indonesian youth movement much later when making the pledge to recognize Indonesia as one motherland, one nation, with one unifying language they called Bahasa Indonesia. Said pledge was proclaimed in October 28, 1928, known as Youth Pledge (*Sumpah Pemuda*).

Below is the original text of the Youth Pledge, written in the then-orthography that was still much influenced by Dutch (“Youth Pledge”, 2021):

¹ Roman number (30)

² The population of Indonesia has always been concentrated in Java Island, with currently more than 140 million habitants, more than half of the country's population. This currently makes Javanese the most populous island in the world (“List of islands by population”, 2022). Despite being populous, consisting of many elites, and having numerous of the historical occurrences there, the choice of unifying language fell to the standardized version of Malay, rather than Javanese, due to its simplicity. In comparison to Malay, Javanese is a language of complex phonology, various registers of speech and varying vocabulary across the island, making it unrealistic as a language to be spoken across the archipelago that consists of thousands of islands and hundreds of local languages (“Indonesian language”, 2022). Another reason was because they wanted to choose a relatively neutral language without favoring one ethnicity over the rest (“Javanese language”, 2022).

Javanese language (2022, January 26). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Javanese_language&oldid=1067983437.

List of islands by population (2022, January 13). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=List_of_islands_by_population&oldid=1065505979.

Indonesian language (2022, February 14). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Indonesian_language&oldid=1070887199.

Pertama

Kami poetra dan poetri Indonesia, mengakoe bertoempah darah jang satoe, tanah air Indonesia.

Kedoea

Kami poetra dan poetri Indonesia, mengakoe berbangsa jang satoe, bangsa Indonesia.

Ketiga

Kami poetra dan poetri Indonesia, mendjoendjoeng bahasa persatoean, bahasa Indonesia.

This text is translated into:

Firstly

We the sons and daughters of Indonesia, acknowledge one motherland, Indonesia.

Secondly

We the sons and daughters of Indonesia, acknowledge to be of one nation, the nation of Indonesia.

Thirdly

We the sons and daughters of Indonesia, uphold the language of unity, Indonesian.

Despite some initial displeasure from Javanese nationalists upon the choice of a language that comes from Malay over Javanese, they finally agreed on using Bahasa Indonesia as unifying language, considering its history as lingua franca during the commerce between the different ethnic groups and with merchants of other nationalities³. Thus, Indonesian became the official language of the country and is used across the archipelago until date, both as L1 and L2.

Indonesian as a language is heavily influenced by other languages that have had contact in different points of the history of the country. The first recorded contact was with Sanskrit from the 7th century. Then, along came other languages of India. But also, through commerce, religious influences and colonization, contacts came from Chinese, Arabic, Persian, Portuguese, Dutch, English, as well as other languages of the archipelago. (Haspelmath & Tadmor, 2009, pp. 688-693)

The numerous foreign contacts have resulted in a rich vocabulary that still has many aspects to be studied and discovered. Through this study, we would like to analyze one small part of this language, more specifically, a study on the influence of Portuguese in the language. It's the Portuguese that was left behind since the 16th century, that is known to have left an impact in various fields of the language.

³ Ardanaeswari, I. (2019, June 25) 25 Juni 1938, Sejarah Kongres Bahasa Indonesia I: Meresmikan Bahasa Persatuan. *Tirto.id*. <https://tirto.id/sejarah-kongres-bahasa-indonesia-i-meresmikan-bahasa-persatuan-ecjH>

And yet, we Indonesians still only know very little of it in comparison to other language influences, so we feel the need to dedicate more time and resources in this study.

0.1. Objective of the Study

We have discovered that scholars, Indonesian and otherwise, have tried figuring out the Portuguese linguistic influence in Indonesian, but so far, all they have produced were glossaries. Even the few dictionaries on loanwords such as Russel Jones' **Loan-Words in Indonesian and Malay** (2007) don't provide much information on the words aside from the specifications of the donor language. Furthermore, at the moment, KBBI Daring, which can be accessed through <https://kbbi.kemdikbud.go.id/Beranda> as the online and most updated version of the standard Indonesian language dictionary does not provide any etymological information on Portuguese loanwords, so we still see the need to make an effort in studying Portuguese loanwords in Indonesian language for the sake of providing further information for users.

In addition, in the second semester of EMLex, we had a module called Dictionary Planning and Production, held by Dr. Dominik Brückner, who also became one of the supervisors of this thesis. This module borne an assignment about planning and producing a multi-volume dictionary on loanwords in Indonesian, with each volume corresponding to one specific donor language. It was that assignment that became the inspiration and foundation for this thesis.

As a long-term goal, we would like to create a dictionary of loanwords in Indonesian, specifically and starting with Portuguese, that has a proper microstructure and etymological information. However, as a preliminary study, we first need to analyze the words in order to later be able to provide a list of dictionary articles that are truly valid and accurate. This factor was what drove us into conducting this study of analyzing Portuguese loanwords in Indonesian language.

0.2. Research Question

Based on the preliminary readings and the current situation of Portuguese loanwords in Indonesian language, we have decided to conduct our study based on the following research questions:

1. Identify the Portuguese loanwords frequently found within the Indonesian language.
2. Why are most of the high-frequency Portuguese loanwords nouns?

3. Do Portuguese loanwords in Indonesian language carry any connotation?
4. Which semantic fields do the high-frequency Portuguese loanwords belong to?

0.3. Structure of the Study

This study will be divided into four chapters: Chapter 1 that will discuss the theoretical background, Chapter 2 that will discuss the methodology, Chapter 3 that will analyze the results of the study and finally, Chapter 4 that shows a representation of the loanwords in a dictionary.

In Chapter 1, we will first discuss the theoretical background of language borrowing in 1.1, as the base of the analysis that will follow. From there, we will enter subchapter 1.2 that dwells more into loanwords in Indonesian in general, followed by Portuguese loanwords, especially, in 1.3. We will see the studies that have been done until date and which novelty we can offer from there.

In Chapter 2, we will show the data source (both online and offline) that we will use for our study. In 2.1, we will explain each online resource, as well as their advantages and disadvantages. Subsequently, we will also explain our methodology for the study in 2.2, which will be explained and analyzed in detail in the following chapter.

Thus, Chapter 3 will contain the results and discussion of our study on Portuguese loanwords in Indonesian language. The chapter will be divided into several subchapters. In 3.1, we will first look into the most frequent Indonesian words and whether any of them are of Portuguese origin, then in 3.2. we will see from another point of view and dwell with the most frequent Portuguese loanwords. Afterwards, in 3.3 we will also look into whether any of the Portuguese loanwords carry any particular connotation and in 3.4, we will divide those loanwords into different semantic fields and analyze them.

Last but not least, in Chapter 4 we will show a representation of our proposed dictionary of Portuguese loanwords. Firstly, in 4.1. we will show a representation that was proposed for an assignment prior to this study and then in 4.2. we will redefine our dictionary representation based on the information we have obtained through this study.

Chapter 1: Theoretical Background

1.1. Language Borrowing

When there is a prolonged contact between two different cultures that have languages that differ from one another, there is a high chance that words or terms would transfer from one language to another. Linguistically speaking, one language would replicate a linguistic feature of another language, either partly or wholly (Durkin, 2014, p. 3). This is called language borrowing, although despite being called such, it does not imply that words are borrowed or loaned with the intention of being given back at a later point. These words would be integrated and remain in the new language. These borrowed words are also called loanwords.

There are two different situations of language transfer, namely imitation and adaptation. Imitation, or borrowing, refers to the situation in which the recipient language does not have the vocabulary used by the donor language and tries its best in imitating and reproducing it. This is usually a phonological case, in which the recipient hears something from the donor and thus tries to imitate and reproduce said word the best they can. However, if the recipient language does not have the sound heard in the recipient language, they are forced to adapt or impose to their own phonological system. Thus comes the process of adaptation, or imposition (Van Coetsem, 1988, p. 8).

Furthermore, in the process of language transfer, there is a donor or source language, as well as the receiver or borrowing language. Usually, it occurs in a situation where the donor holds a higher power or position than the receiver. It may be socially, culturally, politically, or other.

There is usually a stratum (or strate) when a language transfer occurs. When talking about a situation in which one language holds a higher stratum, that language becomes the superstratum, while the other language with the lower stratum becomes the substratum. According to Van Coetsem, “[b]oth substratum and superstratum presuppose the linguistic dominance of the s[ource]l[anguage].” (1988, p. 78) In the case of Dutch colonization in Indonesia, for example, Dutch was considered the language of higher stratum, while the then-Indonesian was considered of lower stratum. Hence, in that case, Dutch was the superstratum, while then-Indonesian was the substratum. This occurrence is also the case with many other colonies (“Stratum (linguistics)”, 2022).

Nowadays, with the globalization and the history of humans discovering, conquering and migrating to other places, loanwords have become a normal part of a language’s vocabulary. According to the

Loanword Typology Project (WOLT) conducted by Haspelmath and Tadmor based on 41 languages of the world⁴, “[n]o language in the sample—and probably no language in the world—is entirely devoid of loanwords”. This means that practically every language in the world has been influenced by others, one way or another.

Language borrowing is different from code-switching. In language borrowing, words are adapted into a language with some changes that may occur, be it in form or meaning, while code-switching is alternating more than one language together in one or more sentences in a conversation, depending on the situation. This mostly happens in bilingual or multilingual communities. (United Language Group, n.d.)

1.2. Loanwords in Indonesian Language

Historically being a commerce hub for centuries between merchants coming from different countries that speak different language, obviously after some time, words would be integrated into the local speech of Indonesian.

In the modern-day Indonesian, it is known that a part of the vocabulary comes from a different language, despite not everyone being familiar with its percentage and history, or where exactly which word comes from. But such is the influence, that the Indonesian artist/writer Alif Danya Munsyi saw the need to write the book titled *9 dari 10 Kata Bahasa Indonesia adalah Asing*⁵, which translates into *9 out of 10 Words in Indonesian is Foreign*, to emphasize this fact. The aforementioned book also gives some interesting text samples to show how much foreign influence is found within the language, such as the following:

"Meski hari gerimis, setelah sembahyang lohor, para santri mengayuh roda sepedanya ke pasar, disuruh paderi membeli koran dan majalah, tetapi ternyata kiosnya disegel sebab bangkrut, jadi mampirilah semuanya di toko buku yang uniknya malah menyediakan perabotan khusus keluarga yang ditaburkan di meja baca, antara lain teko poselen, peniti emas, lap, setrika listrik, serta kalender berfoto artis idola" (Munsyi, 2003, p. 2)

Afterwards, the writer broke down the words of the text and analyzed each word's sources, to obtain the following result:

⁴ The results of the project can be accessed at <https://wold.cld.org/>

⁵ Munsyi, Alif Danya. (2003). *9 dari 10 Kata Bahasa Indonesia adalah Asing*. Jakarta: Kepustakaan Populer Gramedia.

Meski (Portuguese: *mas que*), hari (Sanskrit: title of the God of Sun), setelah (Kawi⁶: *telas*), sembahyang (Sanskrit: *sembah hyang*), lohor (Arab: *dzuhur*), para (Kawi: *para*), santri (Tamil: *santri*), mengayuh (Minangkabau: *kayuh*), roda (Portuguese: *roda*), sepeda (French: *velocipede*), pasar (Persian: *bazar*), disuruh (Kawi: *suruh*), paderi (Spanish: *padre*), membeli (Campa: *ble*), koran (Dutch: *krant*), majalah (Arabic: *majalla*), tetapi (Sanskrit: *tad-api*), ternyata (Javanese: *nyata*), kiosnya (English: *kiosk*), disegel (Dutch: *zege*), sebab (Arabic: *ababun*), bangkrut (Italian: *bancarotto*), jadi (Sanskrit: *jati*), mampirlah (Javanese: *mampir*), semuanya (Sanskrit: *samuha*), toko (Chinese: *to-ko*), buku (Dutch: *boek*), yang (Austronesian: *ia + ang*), uniknya (French: *unique*), malah (Javanese: *malah*), menyediakan (Sanskrit: *sedyā*), perabotan (Betawi: *perabot*), khusus (Arabic: *khusus*), keluarga (Sanskrit: *kulabwarga*), ditaburkan (Hebrew: *tabbwur*), di meja (Portuguese: *mesa*), baca (Sanskrit: *waca*), antara (Sanskrit: *antara*), lain (Kawi: *liyan*), teko (Chinese: *te-ko*), porselen (English: *porcelain*), peniti (Portuguese: *alfinete*), emas (Sanskrit: *amasha*), lap (Dutch: *lap*), setrika (Dutch: *strijkezer*) listrik (Dutch: *elektrisch*), kalender (Dutch: *kalender*), berfoto (Greek: *photos*), artis (English: *artist*), idola (Greek: *eidolon*) (Munysi, 2003, p. 3)

Surprisingly or not, 49 out of 57 words in the text are loanwords from the most diverse sources, making it 85,96% of the total number. It makes complete sense if the writer expressed that nine out of ten words in Indonesian are foreign, or more specifically, have foreign origins.

Sometimes it is not easy to determine what the language donor of a word is. For example, in the text above we see a couple of Greek loanwords, *foto* and *idola*. While it is true that those words initially come from Greek, they actually entered Indonesian through Dutch as the words *foto* and *idool*, respectively. Therefore, based on this fact, rather than stating the etimologia remota Greek, it would be more common to state the etimologia prossima for those words: Dutch. This method is also theoretically more favored by etymologists in general (Buchi, 2015).

As seen in the example above, one of the language donors in Indonesian was Portuguese. In history classes in school, we were taught that the Portuguese arrived in the Indonesian archipelago to trade, spread their religion and conquer our lands, and as a consequence, they were considered our first colonizers⁷.

⁶ Kawi is an old language that used to be spoken in the eastern part of Java, Bali, and Lombok, which was largely influenced by Sanskrit. Nowadays the language is considered archaic and only used in ceremonies and art performances, such as shadow puppet shows ("Kawi language", 2022). Kawi language (2022, January 22). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Kawi_language&oldid=1061636037

⁷ In school we were always taught that we were colonized by the Portuguese, the Dutch and the Japanese, in that order, until we proclaimed our independence on August 17, 1945.

1.3. Portuguese Loanwords in Indonesian and Malay Language

Ever since the middle of the previous century, scholars have tried studying the linguistic influences Portuguese has had on Indonesian, each approaching from different routes. Some try looking into the modern-day Indonesian, some others look from where Portuguese influence was most prominent in East Indonesia and others look further back into Malay of the 16th century, the lingua franca at the time the Portuguese arrived in the archipelago.

Portugal acknowledged the independence of Indonesia in 1949 and opened a representative office in 1950. However, between the years 1965 and 1999, the diplomatic relations between the two countries were broken due to various factors. This factor contributed to the lack of knowledge and attention to Portugal and how the country had culturally influenced Indonesia in the past.

Nevertheless, during that time of unstable diplomatic relations, Portuguese representatives still occupied the office as consul despite remaining without an ambassador. One remarkable Portuguese man called António Pinto da França was on this seat from 1965 to 1970. During his seat, he made a journey around the archipelago and discovered traces of Portuguese culture that ran deeper than what he had thought, thus he manually collected Portuguese-sounding words he heard from locals and compiled them into one glossary, which was included in the book he first published in 1970 (Pinto da França, 2014, pp. 27-35). This glossary becomes one of our starting points at analyzing Portuguese loanwords.

In his book, Pinto da França mentioned having met a historian named Paramitha Abdurachman, who was very engaged in tracing the Portuguese heritage herself. One of her works was also to compile Portuguese loanwords, but based on the vernacular language in Central Moluccas in East Indonesia, which was the area where the Portuguese were mostly based centuries ago. She started conducting her project in 1965 by interviewing locals of different age groups between the ages 15 and 70. By the time her project was done in 1970, she was able to obtain a wordlist, that was also included in her book, alongside some explanations (Abdurachman, 2008, pp. 162-182).

A few years later, in 1973, together with Dr. Denys Lombard and Prof. Luigi Santa Maria, Dr. Russel Jones founded the Indonesian Etymological Project. They were later joined by other scholars of different nationalities to compile loanwords in Indonesian (and Malay) of different language sources, namely Sanskrit, Arabic, Persian, Hindi, Tamil, Chinese, European languages and Japanese. This project borne a dictionary of loanwords with basic etymological information in 2008 (Jones, 2008).

In 2004, the scholars Martin Haspelmath and Uri Tadmor set up a project that aimed in collecting information on loanwords based on sample languages, which they named The World Typology Project. With the help of a group of contributors, they chose 41 representative languages of the world of varying number of speakers and ease of reach (some languages are quite remote), with a sample of 1,000-1,400 words each. This project, funded by the Max Planck Institute for Evolutionary Anthropology in Germany, was conducted until 2008, with the end product a database called The World Loanword Database and can be both accessed online and found in printed version. One of the chosen languages of the project was Indonesian and the database included some Portuguese loanwords as well (Haspelmath & Tadmor, 2009). Additionally, in the book that accompanied their project, Tadmor who did the research for Indonesian identified four major semantic fields that contained loanwords in Indonesian: Religion and belief (70% borrowed vocabulary), the Modern world (66,4%), Clothing and grooming (55,6%) and Law (51,4%) (Haspelmath & Tadmor, 2009, p. 700).

Meanwhile, the Malaysian scholar Ahmat Adam traced his research on Portuguese language influence back to the classical Malay of the 16th century that used to be the lingua franca of both the Malaccan Peninsula and the Indonesian Archipelago. He discovered that classical texts of *Sejarah Melayu* (Malay Annals) written in 1612 or earlier already included some Portuguese vocabulary connected to warfare and navigation. For example, one excerpt mentioned the arrival of Afonso de Albuquerque in Malacca with his fleet as the following:

“Fongso Dhalberkarki pun naik ke Portugal menghadap Raja Portugal minta irmada, maka diberi oleh Raja Portugal empat buah kapal dan lima buah ghali panjang. Maka ia turun berlengkap pula di Guha tiga buah kapal, lapan ghalias, empat buah ghali panjang, lima belas fusta, maka jadi empat puluh semuanya...” (Raffles 18: narrative 23, f.191)

The text above was translated in his book into:

“Afonso Dalboquerque then headed for Portugal to see the King of Portugal in order to request armadas which the King of Portugal then gave him, viz. four ships and five long galleys. He then came down to Goa fully armed with three ships, eight galleys, four long galleons, [and] 15 pinnaces making it 40 altogether...” (Raffles 18: narrative 23, f.191)

We underlined words in the text that were clearly Portuguese influence, namely *irmada* (PT: *armada*), *ghalias* (PT: *galeaca*), *ghali* (PT: *gale*) and *fusta* (PT: *fusta*), which are all connected to navigation, explaining Albuquerque’s fleet and the types of ships he had brought with him. And to conclude, this

book also included a glossary of Portuguese loanwords in Malay, as well as a smaller glossary of Malay loanwords in Portuguese (Adam, 2020, pp. 61-74).

Most recently, the Indonesian scholars Daya Negri Wijaya, Deny Yudo Wahyudi and Siti Zainatul Umaroh conducted a research in Ambon to trace the Portuguese influence in their lands that borne a book. One of the results of their research was that not only was Portuguese a big influence on Ambon Malay's language, but also there were many toponyms in the area with Portuguese-sounding names, such as *Lorong da Silva* (Da Silva Corridor) where the descendants of the Da Silva family still live, *Sungai Olifante* (Olifante River) that is probably adapted from the Portuguese word *elefante* (elephant) and *Batu Capeo* (Chapeo Rock), which likely comes from the Portuguese word *chapéu* (hat), because of its similarity to a hat. The book's chapter on language influence was also complemented with glossaries of toponyms in Ambon and Portuguese loanwords in Malay Ambon. (Wijaya, Wahyudi, Umaroh, 2021, pp. 44-61)

Chapter 2: Methodology

2.1. Data Source

2.1.1. Loanword Online Resources

a. The Corpora Collection of the University of Leipzig

In order to conduct the assessment and analysis of Portuguese loanwords, we will be using the help of Corpora Collection of University of Leipzig⁹. In this website, four different Indonesian corpora are available:


- Mixed Corpus 2013
- News 2019
- News 2020
- Wikipedia 2021

We will be using the Mixed Corpus 2013 for the study, because it contains a mixed source, meaning that it covers various fields, user backgrounds and degrees of speech. Moreover, it is the largest among the four, with 74,329,815 sentences and 1,206,281,985 tokens, several times larger than the other three corpora.

The Corpora Collection itself is a project done by the University of Leipzig, which compiles corpora in different languages and varying sizes, from 10,000 to 1 million sentences, collected randomly from various sources, such as news sites and blogs. These corpora can be used for various means. They can be downloaded, imported and used on other applications or databases as well.

This website will be used not only due to the availability of the largest Indonesian corpus online we could find, but also because the site has features that will be useful for the analysis, especially to determine the frequency of words. Just to give an example, we will show how the word frequency and other useful features are presented on the website (see Figure 1):

⁹ https://corpora.uni-leipzig.de/en?corpusId=ind_mixed_2013# (retrieved on December 23, 2021)



×
🔍
?

Indonesian mixed corpus based on material from 2013 with 74,329,815 sentences.
 [Change corpus](#)

Word: Natal
 Number of occurrences: 49,690
 Rank: 3,054
 Frequency class: 9

Words with Similar Context: natal | Tahun Baru | Hari Natal | Imlek | Paskah

Examples

- Liburan **Natal** tiba. (xshot.winnerinter.co.id, collected on 30/01/2014)
- Artikel Tentang Wallpaper **Natal** yang ada di belfend. (archive.kaskus.co.id, collected on 14/02/2014)
- Abramovich tak memberikan THR **Natal**. (us.bola.viva.co.id, collected on 02/02/2014)
- Massa ini dengan marah mengultimatum orang-orang kristen itu untuk tidak merayakan **Natal**. (indonesia.faithfreedom.org, collected on 08/05/2012)
- Artikel Tentang Dp **Natal** Buat Bbm yang ada di belfend. (www.belfend.web.id, collected on 13/02/2014)
- Contoh Naskah Drama Remaja **Natal**. (www.proberta.com, collected on 08/05/2012)
- Sebarkan rasa kasih dan damai **Natal** kepada orang-orang yang kita sayang. (inkado.umm.ac.id, collected on 06/02/2014)
- Nah ni mau **Natal** nih.. (b2w-indonesia.or.id, collected on 07/02/2014)
- Membuat omamen hiasan **Natal**. (dmanongga.blog.co.uk, collected on 08/05/2012)
- Artikel Tentang Pantun Puisi **Natal** yang ada di belfend. (origin-www.unileverfoodsolutions.co.id, collected on 15/02/2014)

+10 +25 +100

Words occurring in the Same Sentence

Tahun Baru (76,799), perayaan (52,139), Baru (42,352), Tahun (40,305), belfend (26,447), Hari Natal (23,695), Mandailing Natal (23,384), merayakan (21,754), Artikel (20,001), Tentang (19,511), Mandailing (18,755), Perayaan (15,421), Desember (12,519), menjelang (11,870), Hari (11,568), libur (10,986), 25 Desember (9,916), Kabupaten Mandailing Natal (9,326), Hari Raya Natal (9,178), hari (8,830), liburan (7,974), Kristiani (7,947), Selamat (7,789), Pohon Natal (7,684), gereja (6,202), umat (6,200), Ucapan (5,924), Hari Raya (5,876), tahun (5,361), pohon (5,269), pengamanan (5,259), menyambut (5,069), di (4,681), malam (4,379), 2013 (4,289), Gereja (4,035), Kristen (3,815), natal (3,756), Yesus (3,751), Christmas (3,656), Pohon (3,648), misa (3,616), 25 (3,428), dan (3,368), Raya (3,360), idul (3,262), hadiah (3,207), 2014 (3,186), kado (3,128), Misa (3,107), Madina (2,923), Lebaran (2,918), Tahun baru (2,886), bersama (2,883), Argawan (2,849), Paskah (2,720), Idul Fitri (2,719), Kristus (2,717), raya (2,714), Gambar (2,708)

+10 +25

Neighbour Cooccurrences

Left Right

perayaan (67,613), Mandailing (27,826), merayakan (25,771), Perayaan (22,077), Hari (14,138), libur (11,803), menjelang (10,491), Selamat (9,657), liburan (9,629), pohon (9,601), hari (8,700), Hari Raya (7,544), malam (7,511), Pohon (6,247), Ucapan (5,289), Raya (5,201), menyambut (3,690), hadiah (3,587), pengamanan (3,365), raya (3,312), selamat (3,097), Misa (3,022), kado (3,015), misa (2,743), Malam (2,662), sebelum (2,522), Menjelang (2,010), ibadah (1,934), suasana (1,919), Panitia (1,870), lagu-lagu (1,629), jelang (1,521), saat (1,506), ibadah (1,471), makna (1,393), Libur (1,383), momen (1,254), Jelang (1,177), pesta (1,152), pemak-pemik (1,141), Menyambut (1,104), dekorasi (1,079), Rayakan (1,020), Merayakan (1,014), hiasan (1,003), Lagu (989), Puisi (974), Tema (936), Liburan (932), Gambar (889), Ante (883), acara (864), bingkisan (851), tema (829), kebaktian (799), pesan (783), peringatan (747), sukacita (740), angkutan (711), Kado (694)

+10 +25

Word graph

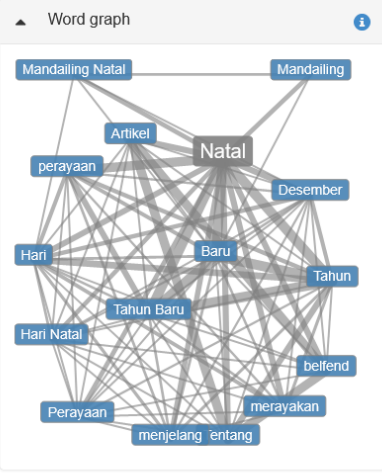


Figure 1 Result page for the entry *Natal* on Corpora Collection of University of Leipzig, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=Natal (retrieved on December 23, 2021)

In Figure 1, we used the Indonesian word *Natal*, loaned from the exact same word in Portuguese *Natal*, which equivalent to *Christmas* in English. On the top of the page, we can see the number of occurrences, rank and frequency class. Based on this information, we can determine that the Indonesian word *Natal* ranks 3,054 on the most frequent words within this corpus.

However useful may the Corpora Collection be, the downside is that it will not include wrong orthography (e.g. typo) or derivative entries, such as affixes, compounds, or the Indonesian possessive pronoun *-nya*. It also detects words at the beginning of a sentence (written with capital letter) as different entries. Moreover, homonyms are not considered as separate cases, should there be any. Thus, the result numbers may deviate from reality, which we will have to keep in mind from the beginning on.

b. Kamus Besar Bahasa Indonesia Daring

The following resource that will become very important for this study will be Kamus Besar Bahasa Indonesia Daring (KBBI Daring). This website is the online version of the standard Indonesian dictionary, which at the moment this study is conducted, features 115,669 entries with the latest update having been in October 2021. Having the most complete number of Indonesian vocabulary, KBBI Daring will allow us to cross-check the meanings and spellings of any of the loanwords that will be analyzed.

In order to show how the website of KBBI Daring works, we will once again use the example word *Natal* and explain the features of said page (see Figure 2):

The screenshot shows the KBBI Daring website interface. At the top, there is a dark blue navigation bar with the KBBI logo, 'KBBI Daring', 'Cari', 'Seputar Laman', 'Halo yuliyudo!', and 'Keluar'. Below the navigation bar, a message reads: 'Halo yuliyudo! Sudahkah Anda mengecek halaman manajemen akun Anda? Anda dapat melihat cara membukanya di sini. Jika Anda pernah mengajukan usulan-usulan, mungkin usulan-usulan tersebut telah diproses oleh redaksi kami.' A search bar contains the word 'Natal'. Below the search bar, the entry for 'na.tal' is displayed. It includes a thesaurus link, two definitions (one in red 'n' for 'kelahiran seseorang' and one in red 'n' for 'kelahiran Isa Almasih'), and a 'Kata Turunan' section with 'natalan'. The footer contains copyright information for Badan Pengembangan dan Pembinaan Bahasa and links for Android and iOS apps.

Figure 2 Result page of the entry Natal on KBBI Daring, from <https://kbbi.kemdikbud.go.id/entri/Natal> (retrieved on December 23, 2021)

The microstructure includes the entry in bold and separated by syllable, followed by a link to the thesaurus entry whenever possible. Next, we have the senses. The red *n* means *nomina* or noun, and sometimes, an example is given after the meaning. Last, there *Kata Turunan* feature or derivative,

which suggests words that have the same root. In this case, *natalan* is a colloquial word that means to celebrate Christmas.

Additionally, as a registered user, we are able to see some extra features such as etymology, which, at the point the study is conducted, has only been developed for Arabic loanwords, while currently the same feature for Sanskrit loanwords are under development. It also has the extra suggestion feature in case the entry search does not give a result, hence intuitively suggesting entry or entries with closest spelling.

c. The World Loanword Database

In 2004, Martin Haspelmath and Uri Tadmor started a project to compile, compare and analyze loanwords from 41 different languages of the world, forming a database of loanwords by the time it finished in 2008, which they then called the Loanword Typology Project (WOLD). This project was helped by a group of contributors and researchers, each expert in the language they were in charge with, with each person working on one language. Tadmor⁹, who had extensive knowledge in Indonesian language himself, contributed for this particular language's sub-database and based in Jakarta in order to do his part.

The project, which can be accessed under <https://wold.cild.org/> , gives a list of small vocabulary sample of each language within the database (1,000-1,400 words each) and we can see not only the meaning of each word, but also a small etymological information that accompanies it.

While the sample may be very small in number compared to how extensive a language is and 41 is a number too small compared to the number of languages in the world as well, it still can give us an idea of the dynamics of the chosen languages as they traveled the world and became donors to other languages.

In order to give an idea of how the result of the project is presented, we will show what the page for Indonesian looks like in the following image (see Figure 3):

⁹ Uri Tadmor was working for the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, at the time this project was conducted. He's currently the Publishing Director at Brill Linguistics (<https://www.linkedin.com/in/uri-tadmor-778b0659/>)



Language Indonesian

Glottocode: Indo1316 ISO 639-3: ind

as recipient as donor

As recipient Language

Indonesian vocabulary contributed by On Yabur

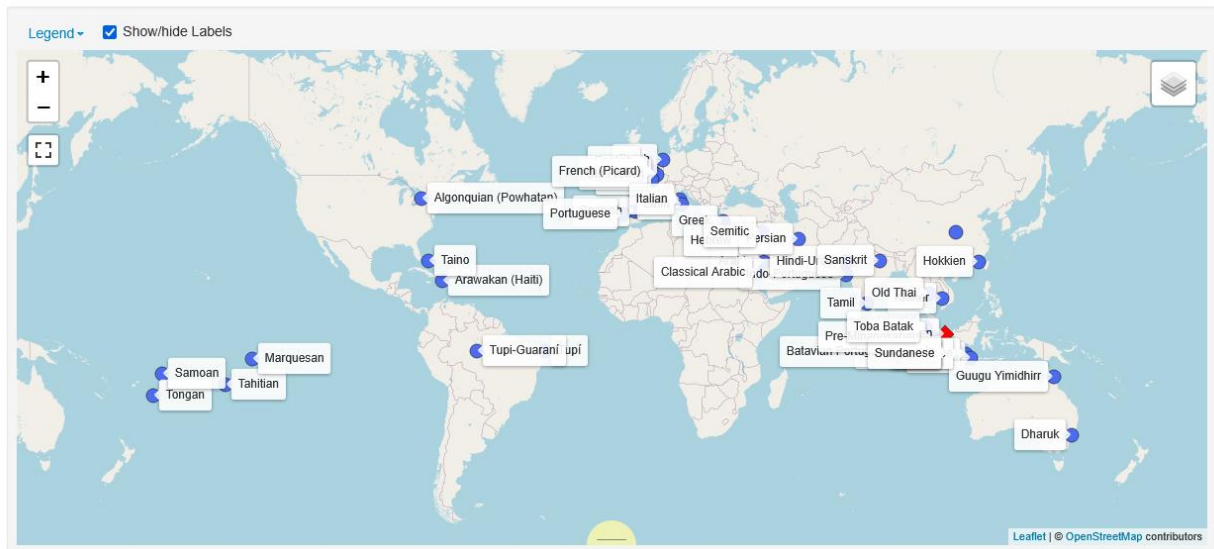


Figure 3 Map of Indonesian as a recipient language and its donor languages on WOLD, from: <https://wold.cld.org/language/27> (retrieved on January 17, 2022)

On Figure 3 above, we can see that Indonesian is recipient of numerous other languages, both near and far, which is curious to see and really makes one wonder how all those languages traveled to the archipelago.

Underneath the map, a list is given with the sample vocabulary of Indonesian. The list can be sorted alphabetically, be it by the entry, the relation (immediate or earlier), its donor language, or the word in the donor language. Said list is shown below (see Figure 4):

Showing 1 to 20 of 1,058 entries

← Previous 1 2 3 4 5 Next →

Borrowed words	Relation	Donor languoid	Source word
<input type="text" value="Search"/>	--any--	<input type="text" value="Search"/>	<input type="text" value="Search"/>
<i>abang</i>	immediate	Unidentified	<i>Unidentifiable</i>
<i>adat</i>	earlier	Arabic	<i>'ādat-</i>
<i>adonan</i>	immediate	Sundanese	<i>adonan</i>
<i>adonan</i>	immediate	Javanese	<i>adonan</i>
<i>agama</i>	earlier	Sanskrit	<i>āgama</i>
<i>aib</i>	earlier	Arabic	<i>'aib</i>
<i>akal</i>	earlier	Arabic	<i>'aql</i>
<i>akhir</i>	earlier	Arabic	<i>āxir</i>
<i>akta kelahiran</i>	immediate	Dutch	<i>akte</i>
<i>akta kelahiran</i>	earlier	French	<i>acte</i>
<i>alam</i>	earlier	Arabic	<i>'ālam</i>
<i>alamat, alamat buruk</i>	earlier	Arabic	<i>'alāmat-</i>
<i>alat</i>	earlier	Arabic	<i>ālat-</i>
<i>altar</i>	immediate	Dutch	<i>altaar</i>
<i>altar</i>	earlier	Latin	<i>altāre</i>
<i>amarah</i>	earlier	Arabic	<i>amārah</i>
<i>ancol</i>	immediate	Sundanese	<i>ancol</i>
<i>andai</i>	immediate	Tamil	<i>aṅṭai</i>
<i>anggur</i>	immediate	Persian	<i>angūr</i>

Showing 1 to 20 of 1,058 entries

← Previous 1 2 3 4 5 Next →

Figure 4 List of loanwords in Indonesian on WOLD, in alphabetical order, from <https://wold.cild.org/language/27> (retrieved on January 17, 2021)

Since we're most interested in Portuguese loanwords for this study, we can type Portuguese in the search box to narrow down the list, and thus, we get the following list (see Figure 5):

Showing 1 to 20 of 33 entries (filtered from 1,058 total entries)

← Previous 1 2 Next →

Borrowed words	Relation	Donor languoid	Source word	
<input type="text" value="Search"/>	--any--	<input type="text" value="portuguese"/>	<input type="text" value="Search"/>	
<i>geréja</i>	earlier	Portuguese	<i>igreja</i>	
<i>tuala</i>	earlier	Portuguese	<i>toalha</i>	
<i>méja</i>	earlier	Portuguese	<i>mesa</i>	
<i>bola</i>	earlier	Portuguese	<i>bola</i>	
<i>serdadu</i>	earlier	Portuguese	<i>soldado</i>	
<i>sepatu</i>	earlier	Portuguese	<i>sapato</i>	
<i>kéju</i>	earlier	Portuguese	<i>queijo</i>	
<i>mentéga</i>	earlier	Portuguese	<i>manteiga</i>	
<i>ténda</i>	earlier	Portuguese	<i>tenda</i>	
<i>jendéla</i>	earlier	Portuguese	<i>janela</i>	
<i>garpu</i>	earlier	Portuguese	<i>garfo</i>	
<i>terigu</i>	earlier	Portuguese	<i>trigo</i>	
<i>keréta</i>	earlier	Portuguese	<i>carreta</i>	
<i>roda</i>	earlier	Portuguese	<i>roda</i>	
<i>minggu, (hari) Minggu</i>	earlier	Portuguese	<i>domingo</i>	
<i>sekolah</i>	earlier	Portuguese	<i>escola</i>	
<i>sumpah palsu</i>	earlier	Portuguese	<i>falso</i>	
<i>berdansa</i>	earlier	Portuguese	<i>dança</i>	
<i>terwélu</i>	earlier	Portuguese	<i>coelho</i>	

Showing 1 to 20 of 33 entries (filtered from 1,058 total entries)

← Previous 1 2 Next →

Figure 5 List of Portuguese loanwords in Indonesian on WOLD, from: <https://wold.cld.org/language/27#recipient-map> (retrieved on January 17, 2022)

The list on Figure 5 now has thirty-three entries of Portuguese, Indo-Portuguese¹⁰ and Batavian Portuguese Creole¹¹. Only four among them are creole and three are Indo-Portuguese, making twenty-six Portuguese loanwords out of the 1,058 words in the Indonesian sub-database.

From this list, we will try checking one entry, in order to see how this project is presented. We will pick one relatively common word, which is *meja*¹² (see Figure 6):

The screenshot shows the WOLD website interface. At the top, there's a navigation bar with 'Home', 'Vocabularies', 'Meanings', 'Languages', and 'Authors'. Below this is a search bar containing 'meja' and a 'cite' button. The main content area displays the entry for 'meja' with the following details:

- Word form:** meja
- LWT meaning(s):** the table
- Analyzability:** unanalyzable
- Age:** Early Malay [27] (500–1700)
- Register:** General

On the right side, the 'Loanword Information' sidebar shows:

- Borrowed status:** 1. clearly borrowed
- Source words:** mesa 'table' Portuguese
- Effect:** No information
- Salience:** No information
- Contact situation:** Portuguese influence [27]

The footer includes a Creative Commons Attribution 3.0 Germany License, the application source (v2009-49-g77272b5), and a GitHub link.

Figure 6 Information page of the entry *meja* on WOLD, from: <https://wold.cld.org/word/71261839983027004> (retrieved on January 10, 2022)

The page shows the contributor, the meaning in English and the approximate time of contact between the two languages, among others. On the right side, it shows some more information on the donor language, which is Portuguese. Also, each / button is clickable and may give some more information on the select field.

Still at the same page on Indonesian, scrolling down to the bottom gives us another map with Indonesian as a donor language, as shown below (see Figure 7):

¹⁰ A group of creoles formed and spoken in the coast of India where there used to be Portuguese Indian settlements, although most of them are either endangered or extinct ("Indo-portuguese creoles", 2022).

Indo-Portuguese creoles (2022, January 19). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Indo-Portuguese_creoles&oldid=1063533786.

¹¹ A recently extinct creole spoken by a small group in Tugu area, Jakarta. It is also called Tugu Creole ("Mardijker Creole", 2022).

Mardijker Creole (2022, January 19). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Mardijker_Creole&oldid=1053076517.

¹² Indonesian is a language that does not use accent marks, however, the words *meja* in this website as seen on Figure 6 is written with one for the sake of distinguishing open e from the closed e (the accent denotes open e) that are normally written the same way, which may cause confusions in pronunciation for those who are not native of the language or not familiar with the word.

As donor Language

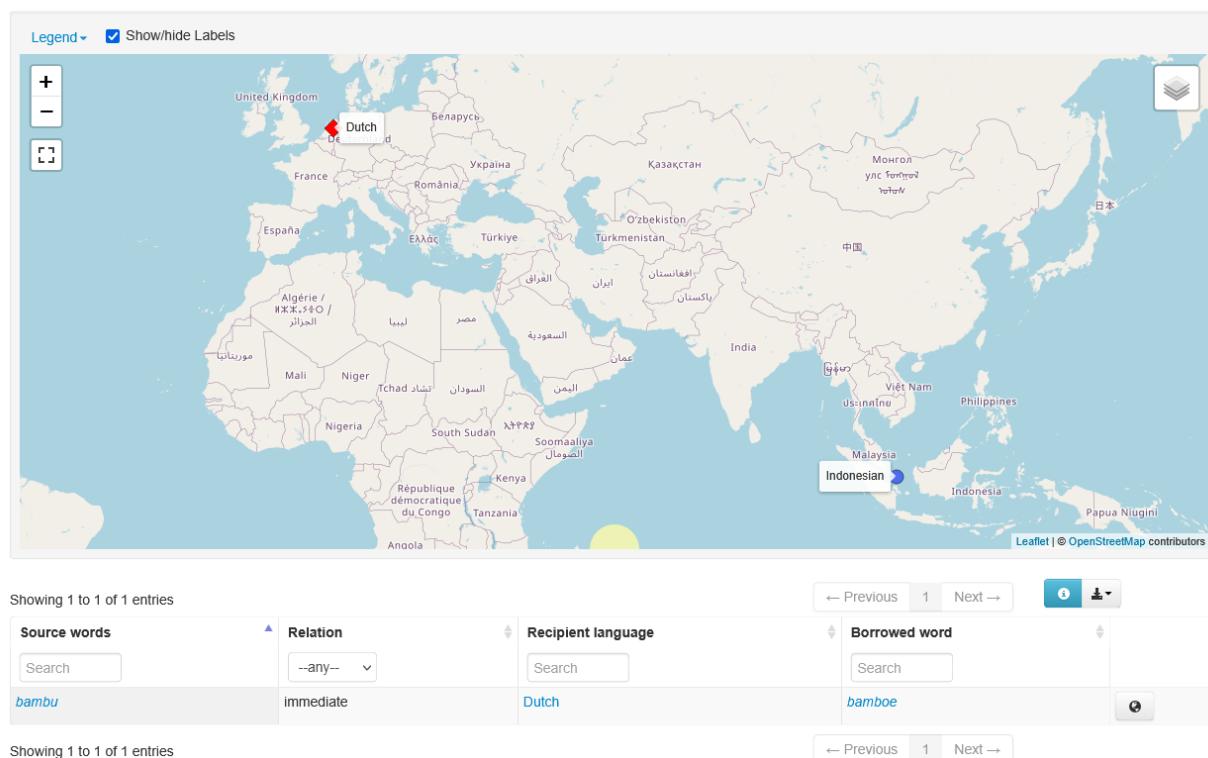


Figure 7 Map of Indonesian as donor language based on WOLD, from: <https://wold.cild.org/language/27> (retrieved on January 17, 2022)

On Figure 7, we see that the language has only been donor to Dutch, at least within the samples among the 41 languages conducted in this project¹³.

Apart from meaning and basic etymological information, this website also has other interesting features, such as semantic fields. There's a given list on fields, where we can click and see words within the entire database that belong to said field. This feature is very useful for comparison purposes.

While this website may not be directly useful for our study in comparison with more complete available glossaries based on the number of samples, it is still resourceful and provides some insights that can be beneficial upon checking the words analyzed during the process, for example if we later on decide to look up for some basic etymological information of certain words or would like to know the semantic field of some.

¹³ English has loanwords from Indonesian as well and are even included in the Oxford dictionary, but despite English being one of the languages of the project, said words were not included in the samples.

2.1.2. Glossaries

Thus far, several authors have gathered and included glossaries of Portuguese loanwords into their works. There are multiple glossaries available, but some are based on the same works, so we only considered the first source if used as reference and compiled everything into one sole list.

The two remarkable works most glossaries are based upon are the books **Bunga Angin Portugis di Nusantara** by Paramitha Abdurachman (2008) and **influência portuguesa na Indonésia** by António Pinto da França (2014), so these will be the major works that will be considered in the making of our list. However, that will not disregard any other addition of vocabulary from other glossaries that we have seen in other works.

2.1.3. Historical and Etymological Dictionaries

In Indonesia, we don't have any dedicated etymological or diachronic dictionary and as mentioned before, even the standard Indonesian dictionary has only started adding etymological information recently. Thus, finding etymological information for Portuguese loanwords especially is a real challenge. Therefore, we will try changing perspectives and instead of looking for the earliest documented information on the language transfer between Portuguese and then-Malay, which may require plenty more resources, we will look up the meaning of the Portuguese source words of those loanwords in historical Portuguese dictionaries from the time of language contact (16th-18th century). Additionally, we will cross-check the etymological information of those words in Portuguese (and Spanish, because of the closeness of the two languages) etymological dictionaries, in order to at least learn the *etimologia remota* from Indonesian, from as many Portuguese sources as possible.

a. **Diccionario da lingua portugueza composto pelo padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva natural do Rio de Janeiro (1789)**

This dictionary written by Rafael Bluteau and perfected by António Moraes da Silva is the closest in timeline and best in content that we were able to find, in relation to Portuguese during the period of the language contact. The dictionary is divided into two volumes and are both available online¹⁴, so they provide an easy access. This work becomes our starting point on obtaining the history of our Portuguese loanwords. It is also known as the first edition of dictionary of Morais.

¹⁴ Silva, António de Morais. (1789) *Diccionario da Lingua Portugueza Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Lisboa, Ferreira. Retrieved March 14, 2022, from <https://digital.bbm.usp.br/handle/bbm/5412> and <https://digital.bbm.usp.br/handle/bbm/5413>

b. Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados (1967)

This etymological dictionary was written by José Pedro Machado and is divided into five volumes. It has etymological information on Portuguese lemmata and gives us some good etymological information on Portuguese words of our concern.

c. Dicionario crítico etimológico castellano e hispánico (1980-1983)

In order to obtain a more thorough etymological information, we will also check the information in this series of dictionary of three volumes by Joan Corominas and José A. Pascual. While this dictionary focuses on Spanish, it still serves us due to the closeness of Spanish and Portuguese.

d. Dicionário Houaiss da língua portuguesa (2003)

Last but not least, we will also look up some etymological information in this dictionary by Antônio Houaiss. While Houaiss was Brazilian and this dictionary in its first release in 2001 was focused on the Brazilian variant of Portuguese, it was re-released in 2003 in three volumes with additional European Portuguese contents. This is a modern-day Portuguese dictionary, but what we will use is the etymology included with each lemma.

2.2. Methodology

First, we will compile glossaries of Portuguese loanwords found in various resources. Thus far, resources that include these glossaries are the following:

- Jumariam, M. T., Qodratillah, & Ruddyanto, C. (Eds.) (1996). *Senarai Kata Serapan dalam Bahasa Indonesia*. Departemen Pendidikan dan Kebudayaan.
- Abdurachman, P. R. (2008). *Bunga Angin Portugis di Nusantara*. Yayasan Pustaka Obor Indonesia
- Wijaya, D. N., Wahyudi, D. Y., & Umaroh, S. Z. (2021). *Nusa Ambon Silang Budaya: Pertemuan Jaringan Global Abad XVI-XVII*. Departemen Pendidikan dan Kebudayaan.
- Jones, R. (general ed.), Grijns, C. D., De Vries, J. W. (Eds.) (2007). *Loan-words in Indonesian and Malay*. Compiled by the Indonesian Etymological Project. KITLV Press.
- Adam, A. (2020). *The Sejarah Melayu Revisited: A Collection of Six Essays*. Strategic Information and Research Development Centre.
- Haspelmath, M. & Tadmor, U. (2009). *Loanwords in the World's Languages*. De Gruyter Mouton.

- Pinto da França, A. & Rogers, R. (2014) *A influência portuguesa na Indonésia: Portuguese influence in Indonesia*. Ed. bilingue em português e inglês. ALIAC-Associação Luso-Indonésia para a Amizade e Cooperação, D.L.

The compilation will sum up into one sole list of known Portuguese loanwords. These loanwords will be listed on a table with the Portuguese source words as well as English equivalents. Then, we will look into the frequency of each word on the Corpora Collection.

Since the study dwells with frequency, we will only choose the entries with highest frequency rank to later analyze. Analyzing the frequency is one important way to study loanwords, “since it is generally assumed that lexical stability increases (and therefore borrowability decreases) with frequency.” (Haspelmath & Tadmor, 2009, p. 15) However, thus far, no research on the frequency of Portuguese loanwords in Indonesian has been made, which drove us to conduct this study. Subsequently, on the latter part of the analysis, we will work on the words’ semantic field, as well.

For comparison purposes, we will also extract the 50 most frequent words in Indonesian from the Corpora Collection, which has this list as one of its features. That way, we can determine whether there are loanwords within the most frequent words and if there are, whether any come from Portuguese.

Furthermore, we would also like to analyze other aspects of the words aside from the frequency per se, such as their semantic fields and connotations. It would be interesting to see which semantic fields those words belong to, and whether any of the loanwords has any negative, positive, or other connotation in Indonesian.

Chapter 3: Results & Discussion

For this study, we will try analyzing the frequency of words in Indonesian seen from different points of view. First, we will see the most frequent words in Indonesian and next, we will determine the frequency of Portuguese loanwords within the Indonesian vocabulary. Afterwards, we will also analyze other aspects of the Portuguese loanwords, namely the connotations and semantic fields.

3.1. Most Frequent Words in Indonesian Language

In this first part of the study, we will pick the 50 most frequent words in Indonesian based on chosen corpus at the Corpora Collection from the University of Leipzig. From that list, we can determine how many of them are loanwords and more specifically, Portuguese loanwords, if any, with the help of the etymology feature of KBBI Online, which, while it is still under development, may help give us an idea. A column of English equivalent is also included for reference.

The 50 most frequent words are the following (see Table 1):

No.	Word	Meaning in English	Loanword	Origin
1.	yang	that (conjunction)		
2.	dan	and		
3.	di	in/on/at		
4.	dengan	with		
5.	ini	this		
6.	untuk	for		
7.	dari	from		
8.	itu	that		
9.	dalam	inside		
10.	tidak	no		
11.	pada	at		
12.	akan	will (denotes future)		
13.	juga	also		
14.	adalah	is/are		
15.	ke	to		
16.	ada	to exist		
17.	bisa	can		
18.	tersebut	said something		

19.	atau	or		
20.	menjadi	to become		
21.	oleh	of		
22.	dapat	can be		
23.	karena	because	✓	Sanskrit
24.	lebih	more		
25.	tahun	year		
26.	sudah	already (indicates past tense)	✓	Sanskrit
27.	sebagai	as		
28.	kita	we		
29.	mereka	they		
30.	Indonesia	Indonesia		
31.	saya	I (pronoun)	✓	Sanskrit
32.	orang	person/human		
33.	telah	already (indicates past tense)		
34.	bahwa	that (conjunction)	✓	Sanskrit
35.	satu	one		
36.	harus	must		
37.	saat	when, moment	✓	Arabic
38.	seperti	like, such as	✓	Sanskrit
39.	kepada	for	✓	Sanskrit
40.	hanya	only		
41.	sangat	very		
42.	masih	still		
43.	lain	other		
44.	secara	as		
45.	banyak	many/much		
46.	kata	word	✓	Sanskrit
47.	para	(indicates plural for person)		
48.	Anda	you (formal)		
49.	merupakan	is/are		
50.	baik	good		

Table 1 The 50 most frequent Indonesian words

The Table 1 above shows that most of the words are conjunction, a case that occurs quite commonly in various languages. We cross-checked the words on the list with the etymology feature at KBBI Daring¹⁵

¹⁵ The etymology feature of KBBI Daring is only provided to registered users. At the moment this study is done, the feature has only been developed for loanwords of Arabic and part of Sanskrit, which are two of the biggest donors of loanwords in Indonesian.

and found that eight of them are known to be loanwords, almost entirely coming from Sanskrit, with one exception from Arabic.

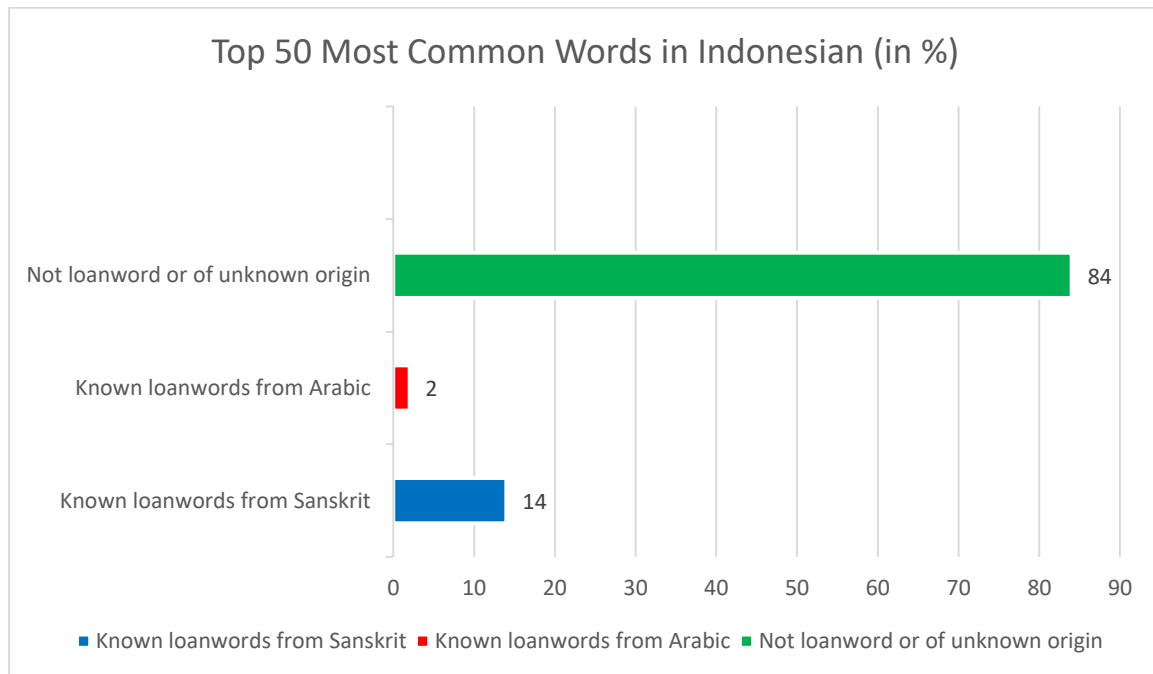


Figure 8 Chart of 50 most frequent words in Indonesian

If we project the list in a chart such as Figure 8 above, we can determine that among the 50 words, 14% among them come from Sanskrit, only 2% come from Arabic, while the rest, or 84% are not known to be loanwords or are of unknown origin.

Since we have thus determined that there are no known Portuguese loanwords on the list and we are not dwelling with loanwords in Indonesian from donors aside from Portuguese, we will leave this information as is at this point and not make any further research on these words, although they will surely be useful for other types of studies.

However, this result leaves us with the questions: how and why are there no Portuguese loanwords among these most frequent words in Indonesian? Where would Portuguese loanwords rank within the Indonesian vocabulary? Also, if we take a look at the list of words on Table 1 above, most of them are conjunctions or particles, with a few pronouns and adverbs. That also raises the question: what are the part of speech of Portuguese loanwords?

3.2. Frequency of Portuguese Loanwords

In this second part, we will see the frequency of loanwords, seen from the Portuguese language point of view, to see how they rank within the Indonesian vocabulary. In order to do that, we take available glossaries from different sources mentioned in the previous chapter and compile them into one sole list. Using that list, we will compare their frequencies using the frequency feature available at the Corpora Collection and create a table of frequencies, as seen below (see Table 2):

No.	Indonesian	Portuguese Origin	English equivalent	Frequency Rank
1.	adven	advento	advent	129,995
2.	algojo	algoz	executioner	30,009
3.	almari (lemari)	armário	wardrobe	34,459
4.	altar	altar	altar	21,750
5.	antero	inteiro	whole	122,973
6.	arena	arena	arena	4,881
7.	arloji	relógio	watch/clock	37,816
8.	armada	armada	armada	5,150
9.	aula	aula	big, multifunctional room	8,841
10.	baluarti	baluarte	stronghold	1,135,748
11.	bangku	banco	bank (to sit on)	4,161
12.	banyo	banjo	banjo	752,441
13.	bantal	avental	pillow	10,682
14.	Belanda	Holanda	Holland	967
15.	baret	bareta	beret	37,000
16.	beludru	veludo	velvet	54,148
17.	bendera	bandeira	flag	3,671
18.	beranda	veranda	veranda	24,906
19.	biola	viola	violin	22,090
20.	bola	bola	ball	918
21.	bolu	bolo	cake	44,015
22.	boneka	boneca	doll	6,584
22.	botol	botelha	bottle	3,463
23.	cerutu	charuto	cigar	40,770
24.	cokelat	chocolate	chocolate	5,558
25.	dadu	dado	dice	30,388
26.	dansa	dançar	dance	28,119
27.	duda	viúva	widower	32,227
28.	gagu	gago	stutterer	139,080
29.	gardu	guarda	guard	17,555

30.	garpu	garfo	fork	29,071
31.	gereja	igreja	church	2,010
32.	gitar	guitarra	guitar	6,080
33.	Inggris	inglês	English	458
34.	jendela	janela	window	3,570
35.	jurnal	jornal	journal	4,085
36.	kaldera	caldeira	volcano crater	53,988
37.	kaldu	caldo	broth	24,986
38.	kamar	câmara	room/chamber	406
39.	kampung	campo	countryside/village	1,908
40.	kanon	canon	cannon	41,851
41.	kantin	cantina	canteen	11,556
42.	kaptan	capitão	captain	5,553
43.	karambol	carambola	fiddle	133,956
44.	kartu	cartão	card	1,105
45.	kastel	castelo	castle	34,671
46.	keju	queijo	cheese	10,044
47.	kemeja	camisa	shirt	10,848
48.	komedi	comédia	comedy	12,140
49.	kompeni	companhia	company	55,042
50.	kereta	carreta	cart/train	1,570
51.	koral	coral	coral	35,382
52.	kubis	couves	cabbage	28,291
53.	kursus	cursos	courses	3,864
54.	kutang	cotão	undershirt	88,104
55.	labu	nabo	pumpkin	19,354
56.	lampion	lampião	lantern	27,909
57.	legenda	legenda	legend	7,518
58.	lelang	leilão	bidding	3,513
59.	lemon/limau	limão	lemon	16,041
60.	lentera	lanterna	lantern	42,043
61.	lenso	lenço	tissue	270,556
62.	loji	loja	shop	92,979
63.	mandor	mandador	foreman	30,480
64.	marakas	maracas	maracas	313,425
65.	markisa	maracujá	passionfruit	49,689
66.	marmot	marmota	guinea pig	145,477
67.	artil	martelo	hammer	60,870
68.	martir	mártir	martyr	32,599
69.	meja	mesa	table/desk	1,982
70.	mentega	manteiga	butter	17,287

71.	meski (pun)	mas que	though	956
72.	Minggu	domingo	Sunday	790
73.	misa	missa	misa	17,397
74.	nanas	ananás	pineapple	16,371
75.	Natal	Natal	Christmas	3,054
76.	nyonya	senhora	missus, lady	3,054
77.	noda	nódoa	stain	11,565
78.	nona	dona	miss	26,212
79.	padri	padre	priest	198,327
80.	palsu	falso	false, fake	2,995
81.	Paskah	Páscoa	Easter	10,100
82.	pastel	pastel	pie	33,061
83.	pegang	pegar	to hold	11,962
84.	peluru	pelouro	bullet	6,319
85.	pena	pena	pen	16,142
86.	peniti	alfinete	safety pin	57,556
87.	pepaya	papaia	papaya	12,057
88.	persero	parceiro	partner, stakeholder	4,244
89.	pesiar	passerar	Cruise	15,452
90.	pesta	festa	party	2,613
91.	Prancis	francês	french	3,014
92.	prangko	franco	stamp	35,954
93.	pigura	figura	frame	54,226
94.	pion	peão (xadrez)	pawn (chess)	75,110
95.	pita	fita	ribbon	7,415
96.	Portugis	português	Portuguese	11,873
97.	puisi	poesia	poem	3,391
98.	rantai	corrente	chain	5,080
99.	renda	renda	lace	45,389
100.	roda	roda	wheel	2,328
101.	ronda	ronda	night patrol	34,024
102.	rosario	rosario	rosary	53,406
103.	Sabtu	sábado	Saturday	888
104.	sabun	sabão	soap	6,621
105.	saku	saco	pocket	8,669
106.	salto	salto	jump	50,526
107.	seka	secar	to dry	85,957
108.	sekolah	escola	school	215
109.	selada	salada	salad	34,049
110.	selai	geleia	jam	24,568
111.	sepatu	sapato	shoe	2,549

112.	serdadu	soldado	soldier	18,423
113.	silet	gilete	razor	52,765
114.	sinyo	sinhô/senhor	young man of European origin	252,412
115.	sita	citar (confiscar)	To confiscate	22,531
116.	tanjidor	tangedor	a traditional musical style	99,766
117.	tapioka	tapioca	tapioca	32,382
118.	teledor	traidor (negligente)	careless	60,355
119.	tembakau	tabaco	tabacco	5,550
120.	tempo	tempo	time/speed	4,414
121.	tenda	tenda	tent	6,682
122.	terigu	trigo	wheat flour	10,521
123.	tinta	tinta	ink	9,549
124.	tolol	tolo	stupid	33,860
125.	tuala	toalha	towel	84,433
126.	tukar	trocar	to exchange	3,061

Table 2 Frequency of Portuguese Loanwords in Indonesian

While we are aware of the size of Table 2 that may exceed the usual table size in the body of a thesis work, we decided to still display here the entire list (and display it once again in Appendix 1) because the content is integral to the study and can neither be summarized nor cut down.

As mentioned in Chapter 2 about Corpora Collection not detecting homonyms as separate words, we also decided to check on KBBI Daring for any possible homonyms out of the 126 Portuguese loanwords. Upon checking, we found out that indeed, there were a number of homonyms, which we will list below with their meanings:

1. baret: a. scratch
 b. beret
2. bolu: a. cake
 b. milkfish
3. cokelat: a. chocolate (tree or sweets)
 b. brown
4. dadu: a. dice
 b. pink
5. kamar: a. room, chamber

- b. (arc.) moon
6. kanon: a. land tax that goes on for generations
b. canon (music)
c. cannon
d. canon (literature)
e. dogma
7. koral: a. coral
b. room or enclosed space to put the stock in cage
8. labu: a. pumpkin
b. necked glass laboratory bottle
c. a type of drum made of leather from Roti Island
9. legenda: a. legend (historical) or someone very famous
b. legend (explanation in maps, pictures, etc.)
10. lenso: a. a type of couple dance from the Moluccas that uses handkerchief
b. (reg.) handkerchief¹⁶
11. nona: a. missus
b. custard apple, wild soursop
12. pastel: a. savory pastry shaped like a half moon, filled with meat and others
b. pastel (color)
13. pita: a. ribbon
b. band (electronics)
14. sabun: a. soap
b. white or yellowish white
15. saku: a. pocket
b. (arc.) to separate
16. seka: a. to wipe
b. an ethnic group living between Indonesian Papua and Papua New Guinea
17. sepatu: a. shoe
b. flower
c. overlock foot
18. sita: a. to confiscate

¹⁶ Term used in the Papua Island

- b. (cl.) pure white
19. tempo:
- a. time, timeline
 - b. tempo (music)

Based on the nineteen words and their homonyms above and comparing with the meanings on Table 2, we can determine which meanings come from Portuguese. Moreover, according to the information provided in the microstructure of each entry on KBBI Daring, some of the homonyms are very rarely used, archaic, or come from Classical Malay, so we could say that these as good as never appear in modern sources compiled in Corpora Collection. Nevertheless, some of these affect and deviate the number of occurrences in the corpus. As Philip Durkin said, homonymy “is very difficult to define and identify if language is looked at purely from a synchronic, non-historical viewpoint,”¹⁷ (2014, p. 36) hence, unless we dedicate time and resources to define each meaning and separate the occurrences, it is unavoidable.

How we were able to tell homonyms and polysemy apart were because they are presented differently in KBBI Daring. A polysemy is written under the same lemma, but they are separated by numbered senses, while a homonym is written under a different lemma.

Now, looking back to Table 2, on a quick glance, we can see that Portuguese loanwords do not belong to the very highly frequent words within the language, as the numbers are relatively big. However, we can still narrow the list down to a certain percentage and check those words based on KBBI Daring’s entries.

Before taking that step we would like to know first how up to date the KBBI Daring page currently is, considering it is updated on a regular schedule. In order to do that, we went to the statistics and update page, and obtain the following information (see Figure 9):

¹⁷ Using the British National Corpus (BNC) to analyze the frequency of words in English, Durkin made the comparison between *rest* in *to take a rest* and *rest*, which are identical in form, yet very different in meaning and origin. The former word has Germanic origin, while the latter is French in origin (Durkin, 2014, p. 36).

Halaman Statistik

■ Kamus

Entri Dasar:	52,326 (45.24 %)
Kata Turunan:	27,440 (23.72 %)
Gabungan Kata:	32,325 (27.95 %)
Peribahasa:	2,077 (1.80 %)
Idiom:	270 (0.23 %)
Ungkapan:	1,142 (0.99 %)
Varian:	89 (0.08 %)
Entri Total:	115,669 (100.00 %)
Makna Total:	134,858
Contoh Total:	30,598
Kategori Total:	254
Makna Per Entri:	1.166
Contoh Per Makna:	0.227

Figure 9 <https://kbbi.kemdikbud.go.id/Beranda/Statistik> (retrieved on January 12, 2022)

Pemutakhiran

Pemutakhiran KBBI dilakukan dua kali setahun, yaitu pada bulan April dan Oktober. Pemutakhiran KBBI terakhir, dilakukan pada bulan Oktober 2021.

Pemutakhiran periode Oktober 2021 ini ditandai dengan masuknya entri baru yang berupa istilah-istilah dari berbagai bidang ilmu. Di antara bidang-bidang ilmu tersebut adalah tata boga (*aburage, baklava*); perhotelan (*arkade hotel, duvet*); hukum (*delik dolus, hak bezit*); psikologi (*ekofobia, hipomania*); nama negara, bahasa, dan mata uang (*Swedia, Katalan, taka*); nama surah (*Al-Lahab, Maryam*); dan kata-kata umum lain, termasuk kata yang dipakai dalam ragam cakapan (*gokil, kudet*) serta kata tidak baku yang dirujuk ke bentuk baku. Selain itu terdapat juga perbaikan dan penambahan entri yang berhubungan dengan ikan, burung, serangga, dan mamalia dengan menyertakan narasumber dari Pusat Riset Biologi, BRIN. Istilah-istilah yang berhubungan dengan wabah Covid-19, seperti *epidemiolog, badai sitokin*, juga dimasukkan sebagai tambahan entri baru.

Adapun perincian data pemutakhiran KBBI periode Oktober 2021 tersebut adalah sebagai berikut.

Pemutakhiran KBBI Daring Oktober 2021

- **Entri Baru:** 1.005
- **Makna Baru:** 1.144
- **Contoh Baru:** 117
- **Perubahan Entri:** 202
- **Perubahan Makna:** 393
- **Perubahan Contoh:** 7
- **Entri Non Aktif:** 14

Figure 10 <https://kbbi.kemdikbud.go.id/Beranda/Pemutakhiran> (retrieved on January 12, 2022)

The statistics page seen on Figure 9 says that the website has to date 115,669 entries. The update as seen on Figure 10 is conducted twice a year, every April and October, and its last update was in October 2021. There is information on the new entries as well, but it is not relevant for our study.

Now, after having made sure that we know the current total number of entries on the website, we will go to the next step, where we will only pick 10% of the total number, which sums up to 11,556 entries. However, for the sake of easier analysis, we will round the amount down to the 10,000 most frequent words, or 8.65% of the total number of entries. Based on this, we are able to eliminate the less frequent words from Table 1 and obtain the following list of words instead, which is sorted according to frequency rank (see Table 2):

No.	Indonesian	Portuguese Origin	English Equivalent	Frequency Rank
1.	sekolah	escola	school	215
2.	kamar	câmara	room/chamber	406
3.	Inggris	inglês	English	458
4.	Minggu	domingo	Sunday	790
5.	Sabtu	sábado	Saturday	888
6.	bola	bola	ball	918
7.	meski (pun)	mas que	though	956
8.	Belanda	Holanda	Holland	967
9.	kartu	cartão	card	1,105
10.	kereta	carreta	cart/train	1,570
11.	kampung	campo	countryside/village	1,908
12.	meja	mesa	table/desk	1,982
13.	gereja	igreja	church	2,010
14.	roda	roda	wheel	2,328
15.	sepatu	sapato	shoe	2,549
16.	pesta	festa	party	2,613
17.	palsu	falso	false, fake	2,995
18.	Prancis	francês	French	3,014
19.	Natal	Natal	Christmas	3,054
20.	nyonya	senhora	missus	3,054
21.	tukar	trocar	to exchange	3,061
22.	puisi	poesia	poem	3,391
23.	botol	botelha	bottle	3,463
24.	lelang	leilão	bidding	3,513
25.	jendela	janela	window	3,570
26.	bendera	bandeira	flag	3,671
27.	kursus	curros	courses	3,864
28.	jurnal	jornal	journal	4,085
29.	bangku	banco	bank (to sit on)	4,161
30.	persero	parceiro	partner, stakeholder	4,244
31.	tempo	tempo	time/speed	4,414
32.	arena	arena	arena	4,881
33.	rantai	corrente	chain	5,080
34.	armada	armada	armada	5,150
35.	tembakau	tabaco	tobacco	5,550
36.	kapten	capitão	captain	5,553
37.	cokelat	chocolate	chocolate	5,558
38.	gitar	guitarra	guitar	6,080
39.	peluru	pelouro	bullet	6,319
40.	boneka	boneca	doll	6,584

41.	sabun	sabão	soap	6,621
42.	tenda	tenda	tent	6,682
43.	pita	fita	ribbon	7,415
44.	legenda	legenda	legend	7,518
45.	saku	saco	pocket	8,669
46.	aula	aula	big, multifunctional room	8,841
47.	tinta	tinta	ink	9,549

Table 3 Portuguese loanwords among the 10,000 most frequent words based on KBBI Online entries

Thus, we can see on Table 3 that there are forty-seven Portuguese loanwords that rank 10,000 or higher, which make 37.3% of the total loanwords from the previous table (see Table 2). However, due to the difficulty in following each word on a table and for the sake of easier viewing of the frequency ranks, we transformed the table into a chart, as seen in the following Figure 11:

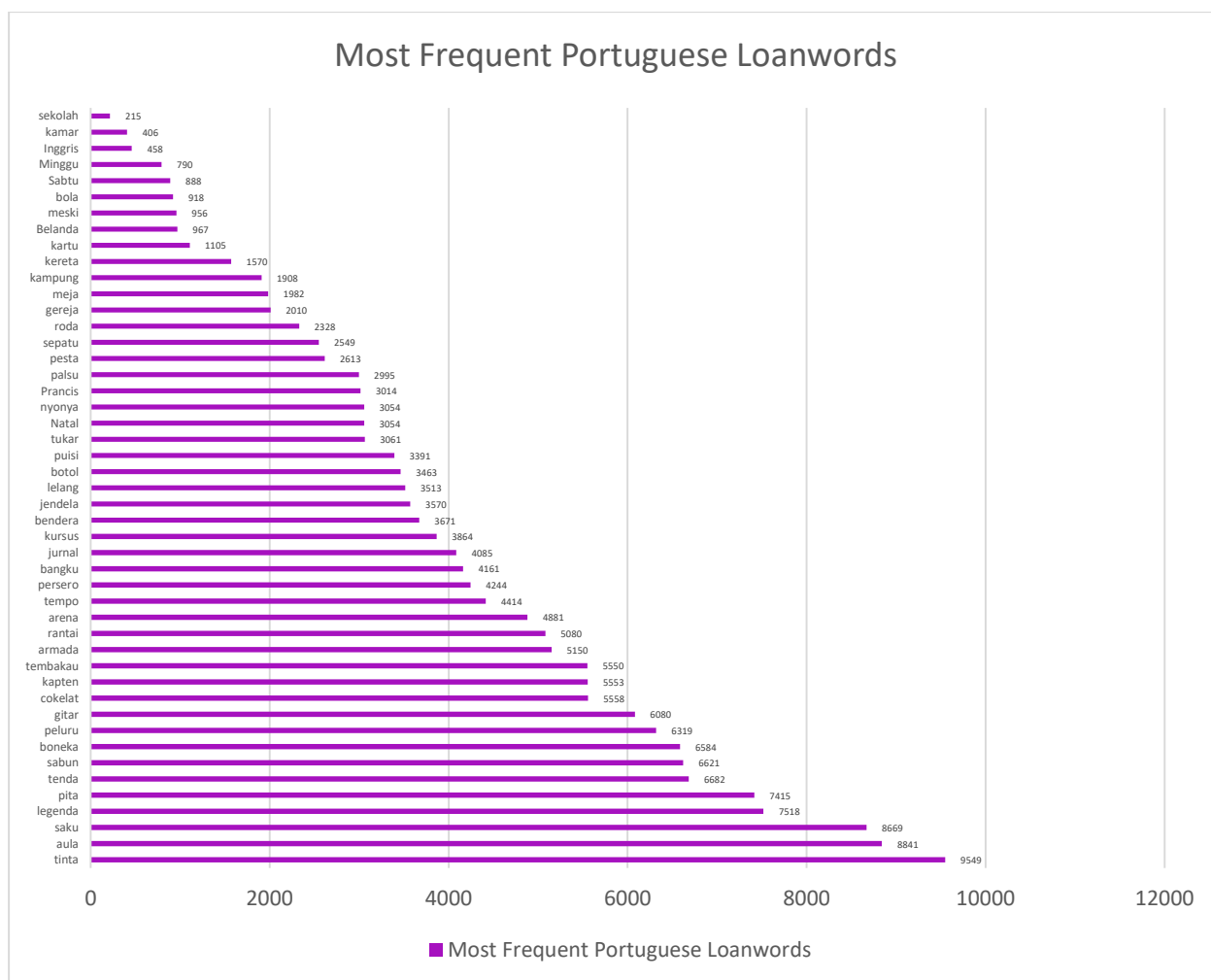


Figure 11 Chart of most frequent Portuguese loanwords

Looking at the chart on Figure 11, which is increasing in number but decreasing in rank, meaning that the lower the frequency and consequently the rank, the bigger the number, we can see both the words and each of their respective ranks more easily. Also, from this new number and view, we can divide the words further by breaking down the information according to the levels of frequency. For this, the categories will be divided into words ranking 1,000 and above, 5,000 and above and 10,000 and above. From this categorization, we obtain the following information (see Figure 12):

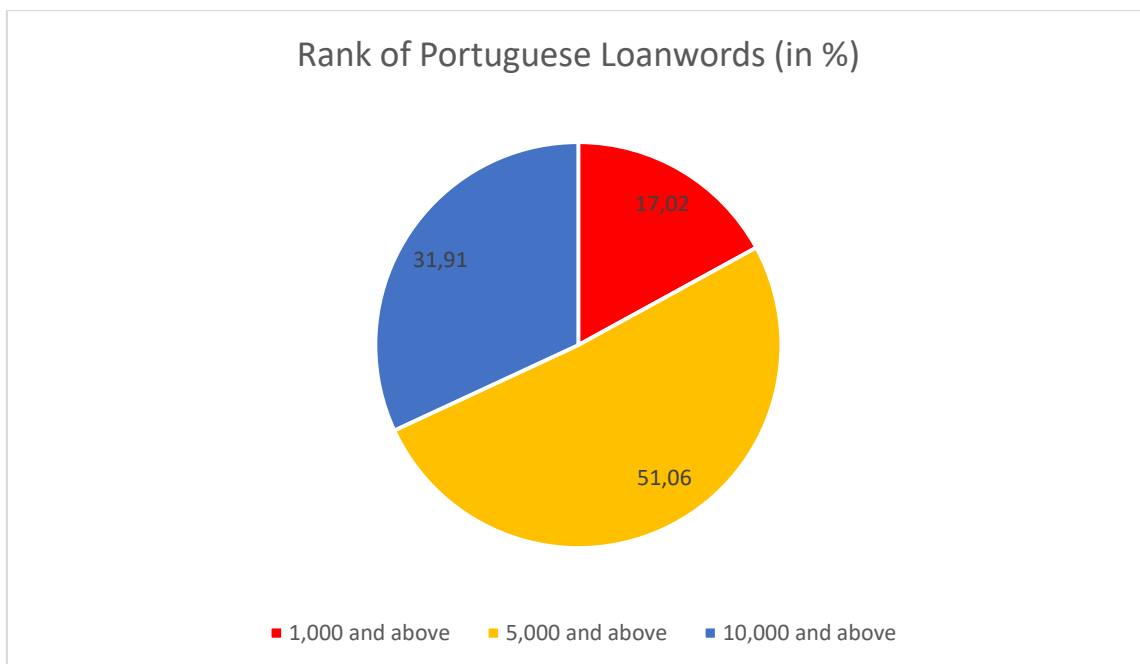


Figure 12 Rank of Portuguese loanwords

With eight or 17.02% of the words ranking 1,000 or above, twenty-four or 51.06% words ranking 5,000 and above, as well as fifteen or 31.91% words ranking 10,000 and above, we can determine that at least half of the forty-four words on the list rank between 1,001 and 5,000.

Now, if we go back to the information on Figure 11 above, the most frequent words (ranking less than 1,000) would be *Belanda* (Holland), *bola* (ball), *Inggris* (England), *kamar* (room), *meski* (though), *Minggu* (Sunday), *Sabtu* (Saturday) and *sekolah* (school).

It is very interesting to find out how the country names *Holland* and *England* would be used more frequently than *Portugal* itself, which ranks only 11,873 and not included in our analysis, despite the loanwords coming from Portuguese. Most probably, it is because as history went by, Dutch colonization

has left far more impact and in more recent years, English has been used much more commonly, causing the words to appear more frequently.

As for *Sabtu* and *Minggu*, they appear very frequently because of being the days of the week and these are words quite commonly used in texts, especially in news. For comparison purposes, we also checked the other days of the week on the Corpora Collection for frequency and obtained the following (see Table 4):

Indonesian	English	Frequency Rank
Senin	Monday	541
Selasa	Tuesday	582
Rabu	Wednesday	549
Kamis	Thursday	583
Jumat	Friday	619
Sabtu	Saturday	888
Minggu	Sunday	790

Table 4 Frequency of Indonesian names of the days of the week

As we can see on Table 4, the days of the week all appear rather frequently in corpora, ranking between 541 and 888, although for a certain reason, the weekend names appear a little less often (ranking 888 and 790, respectively) in texts than the working days (ranking between 541 and 619). But we will not go into further details as it requires further studies to determine the reason.

Moreover, while we were analyzing the days of the week, we determined that aside from *Sabtu* and *Minggu* that are Portuguese loanwords, all the other days are loanwords of Arabic origin. This brought a thought to mind, that historically, Arabic-speaking merchants arrived at the archipelago much earlier with the earliest records being from the 14th century. The connection is very much related to the Islamization of the archipelago, which continues until the present day (Jones, 2008, p. xxiii). Some centuries after, when the Portuguese arrived and started spreading Catholicism in addition to carrying on trade, the words *sábado* and *domingo* that are very related to Catholicism were left behind, together with some other words in the field of religion. *Sábado* was adapted into *Sabtu*, while *domingo* was adapted into *Minggu*.

While on this category, we discovered that *meski* and *palsu* are the only words on the list that are not a noun. In Indonesian, *meski* falls under the particle or connecting word category, while *palsu* is an adjective. If we remember from the previous sub-chapter, we saw that conjunctions or particles have the tendency to rank higher due to its use, so we will conclude that it is the case with this word as well.

3.3. Connotations of Portuguese Loanwords

One thing of interest that caught our attention when we decided to analyze Portuguese loanwords was whether those words have any connotation in particular.

Badan Pengembangan dan Pembinaan Bahasa (Language Development and Fostering Agency) of the Ministry of Education, Culture, Research and Technology, the department in charge of KBBI, has once released an article on how words become part of KBBI¹⁸ and the pre-requisites are: unique, euphonic, adaptable to Indonesian grammar, has no negative connotation, as well as has high frequency and range.

On connotations, the agency states that when there are two words with similar meaning, they will choose the one with no negative connotation, or at least the one with the more positive connotation, with the argument that words with a negative connotation may not be acceptable for higher level users. However, it is not clearly explained what it is meant with higher level. Is it the education level? Is it the status or title? Or are there other possible meanings?

But this fact does not back us down from trying to check whether any of the Portuguese loanwords has a negative connotation. We will try figure it out by using the word graph feature in a word's entry page on Corpora Collection, using the words on the list of most frequent Portuguese loanwords on Table 3.

Historically, the relation between Indonesia and Portugal in languages and cultures have happened in colonial context, which would have emphasized the racism, prejudice, discrimination and glorification of foreigners between the locals and Europeans, as well as in the vocabulary.

The prejudice and deviations registered in the vocabulary use should be included in the dictionaries while at the same time avoiding thoughts of what is politically correct. The use of words that are racist, sexist, xenophobic, etc. obviously need to be labelled as such, as not to contribute its legitimization in the definition and examples used.

¹⁸ Bidang Pengembangan, Pusat Pengembangan dan Perlindungan (n.d.) *Bagaimana Sebuah Kata Masuk ke KBBI*. Retrieved January 21, 2022 from <https://badanbahasa.kemdikbud.go.id/lamanbahasa/artikel/2547/bagaimana-sebuah-kata-masuk-ke-kbbi/>

In this sense, we think it is relevant to identify, wherever possible, every use of the Portuguese loanwords in Indonesian, as well as describe, also quantitatively, how much of these are marked connotatively (positive, negative or neutral).

We start from the hypothesis that the information about how these loanwords are combined (aside from the pragmatic or contextual information, the access of which is more difficult) could deliver some semantic information and its use (a specific collocational use, for example), which helps identify the semantic differences (Iriarte, 2001, pp. 137) and eventually the connotative values.

Upon checking the forty-seven loanwords one-by-one, we determined that there was indeed no word that had a negative connotation, in the sense that the co-occurrences did not give a negative connotation either, although we did see some interesting results. We will show here one representative graph from each level of frequency as we had broken down in the previous sub-chapter.

Firstly, we chose *kamar* (room) for rank 1,000 and above (see Figure 13):

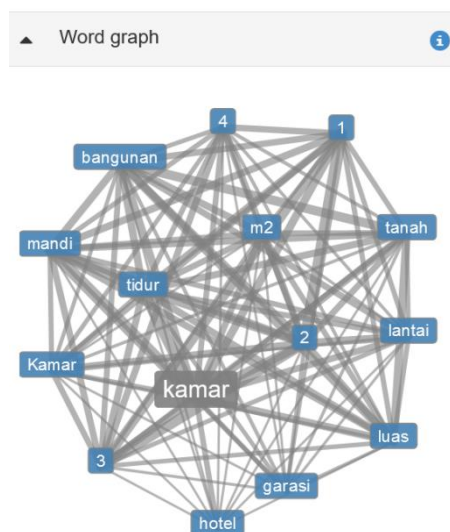


Figure 13 Word graph of the loanword *kamar* (room) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=kamar (retrieved on January 22, 2022)

Most of the co-occurrences on the word graph on Figure 13 are related to a house. We can see that there are numbers between 1 and 4, certainly related to the number of rooms in a house in advertisements or articles, and so does *m2* (or more specifically, *m²*). There is also *hotel* (hotel), making the compound *kamar hotel* (hotel room). However, we were a little intrigued by *garasi* (garage). We

assume it relates to housing as well, but the two words are normally not co-occurrences, because *kamar garasi* (garage room) or *garasi kamar* (room garage) make little sense.

Next, we chose the word *palsu* (false, fake) for the rank 5,000 and above (see Figure 14):

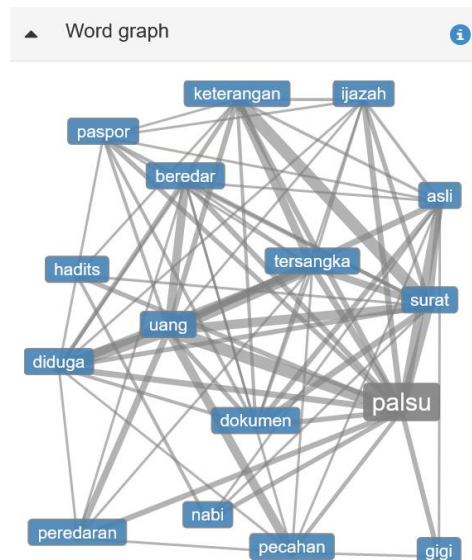


Figure 14 Word graph of the loanword *palsu* (false) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=palsu (retrieved on January 22, 2022)

We chose this word because despite not having a negative connotation on its own, its meaning is not exactly positive either, so that when put together with its co-occurrences, they may create compounds that do not sound really positive. For example, *dokumen palsu* (fake document), *uang palsu* (fake or counterfeit money), or even *nabi palsu* (false prophet), none of which have a very pleasant-sounding meaning, yet it does not mean that the word *palsu* itself has a negative connotation. As seen in the case of *gigi palsu* (false teeth), for example, it does not give the same negative connotation, since it simply means dentures, so we concluded that the connotation depends on the co-occurrences.

Last but not least, from the rank 10,000 and above, we chose the word *peluru* (bullet) (see Figure 15):

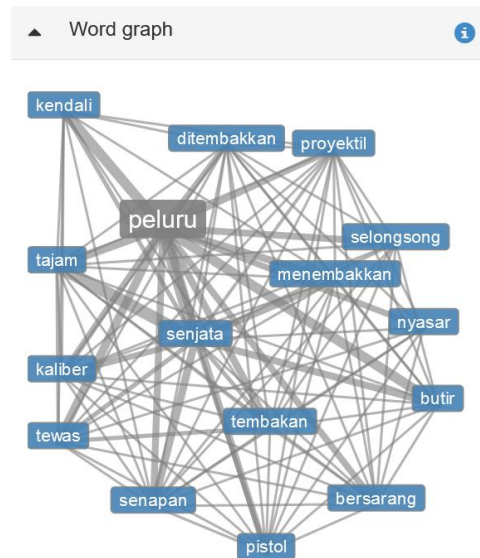


Figure 15 Word graph of the loanword *peluru* (bullet) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=peluru (retrieved on January 22, 2022)

We chose this word with a similar reason than the previous word, *palsu*, on Figure 14. The word *peluru* itself has neither negative meaning nor negative connotation, but the co-occurrences may give one. For example, the compounds *peluru nyasar* (wrongly shot, unfairly targeted), *peluru bersarang* (bullet stuck inside something or someone), or *tewas akibat peluru* (death by bullet shot). We were especially intrigued by the compound *peluru nyasar*, words that often create unpleasant feelings on the listener or reader, because they imply a mistake by the shooter which wounds (even fatally) someone else who was at the wrong place and the wrong time.

3.4. Semantic Fields of Portuguese Loanwords

Next, we analyzed the semantic fields of the most common Portuguese loanwords, in order to know how the language was adapted into then-Malay. For this analysis, we categorized the words into different fields that make sense¹⁹, as follow:

¹⁹ We did not include *meski* into any of the semantic fields, considering it was a stand-alone case of being a particle and would not fall into any of the categories made here.

3.4.1. Commerce

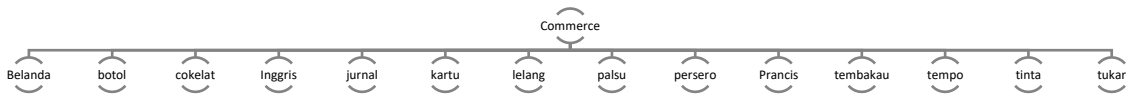


Figure 16 Semantic field of Portuguese loanwords in Indonesian: commerce

There are fourteen Portuguese loanwords that we categorized into the semantic field commerce as seen on Figure 16, which was one of the major activities the Portuguese conducted in the archipelago.

We previously said that *meski* was a no-noun case among the list of words. However, there was actually another word that is not a noun, namely *palsu* (false, fake), which is an adjective. We still included the word into the field of commerce, because it does make sense to put it under said category. After all, it is a word related to trade, such as in occasions of considering and negotiating the authenticity of goods.

In order to show how the word *palsu* is used in modern-day Indonesian and whether it still applies to the semantic field we have assigned the word to, we have selected a couple of example sentences from the corpus used in Corpora Collection:

1. Hasilnya, ditambahkan Kapoltabes, didapat ribuan produk palsu. (www.sinarharapan.co.id, collected on 26/04/2012)
The result, added the Head of City Police, there were thousands of fake products.
2. Ironis, hingga kini masih banyak pemilih yang terkecoh pada janji-janji palsu kandidat. (pt-palangkaraya.go.id, collected on 07/02/2014)
Ironically, until today many voters are still deceived by empty promises of the candidates.
3. Rias mata sebelum memasang bulu mata palsu. (www.thefaceshop.co.id, collected on 03/02/2014)
Apply eye make-up before putting on your false lashes.

As we can see, the word *palsu* now has a wide use, from false, fake, or imitation product (usually something that can be seen and/or touched) to false or empty things that are more abstract (such as *promises* in the example 2). It actually does not deviate much from its donor word *falso* and can really be used in the field of commerce.

Although the chosen name of the semantic field may be commerce, it is also meant to be seen in a wider term. After all, some words are not merely used in that field nowadays, such as the case with *tempo*. And in order to explain that, we will put some sample sentences below:

1. Reaksinya luar biasa, dalam tempo kurang dari sebulan saya sehat kembali dan bisa jalan normal lagi. (seputarkarawang.blogspot.com, collected on 08/05/2012)
The reaction was amazing, in less than a month I was healthy and able to walk normally.
2. Jadi bagaimana bentuk penjajahan tempo dulu? (martinusadyh.web.id, collected on 31/01/2014)
So, what was the colonization like in the old days?
3. Pembayaran harus dilakukan sebelum atau pada tanggal jatuh tempo yang tertera pada Lembar Penagihan. (lpse.pelalawankab.go.id, collected on 30/01/2014)
The payment needs to be made before the due date stated on the billing.

Although translating the sentences into English does not really give the literal equivalence of the word, from these sentences we can more or less see how the words are used in modern-day Indonesian. Since we did not really see the connection with commerce other than the case of the fixed expression *jatuh tempo*²⁰ (due), we tried analyzing the word using its word graph on Corpora Collection (see Figure 17):

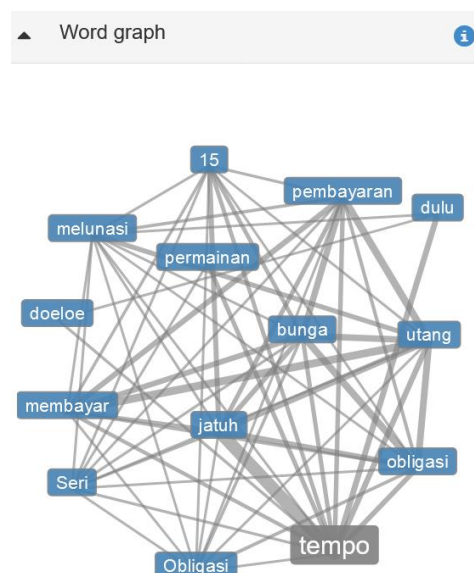


Figure 17 Word graph of the loanword *tempo* (time, speed) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=tempo (retrieved on January 21, 2022)

²⁰ KBBI Daring. (n.d.) Tempo. In KBBI Daring. Retrieved January 31, 2022, from <https://kbbi.kemdikbud.go.id/entri/jatuh%20tempo>

On the word graph on Figure 17 above, we see that the words most related to tempo are *jatuh* (fall), *obligasi* (bonds), *membayar* (to pay), *pembayaran* (payment), *dulu* (beforetime), *bunga* (interest) and *permainan* (game). Therefore, we can conclude that the word indeed has relations with commerce, although mostly in the context of *jatuh tempo*, as mentioned earlier.

3.4.2. Entertainment

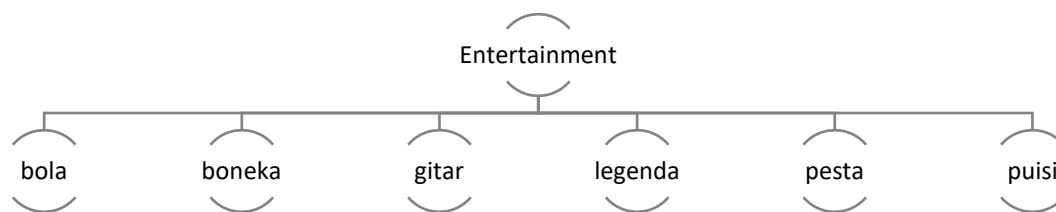


Figure 18 Semantic field of Portuguese loanwords in Indonesian: entertainment

In this semantic field of entertainment as seen on Figure 18, we included six words, namely *bola* (ball), *boneka* (doll), *gitar* (guitar), *legenda* (legend), *pesta* (party) and *puisi* (poem/poetry). These words are quite self-explanatory in this field, except for *legenda*, because as we mentioned in sub-chapter 3.2, *legenda* was one of the homonyms among the loanwords. While both senses equivalent to *legend* in English, it is not the case in Portuguese. In Portuguese, *legenda*²¹ refers to the caption or information in a map/picture, or even subtitle in a video. Meanwhile, to refer to a historical story or someone very famous, the Portuguese equivalence would be *lenda*²².

This fact left us with a bit of confusion, so in order to ensure our hypothesis, we selected a few sample sentences from Corpora Collection:

1. Cerita-cerita ini juga bicara tentang Ramayana serta legenda lainnya. (bem-fh.umm.ac.id, collected on 04/02/2014)
These tales also tell about Ramayana and other legends.
2. Sang legenda juga mendapatkan dukungan dari banyak pihak, kawan atau lawan, salah satunya Lauda. (ns1.kompas.web.id, collected on 02/02/2014)
The legend also received support from friend and foe alike, one of them was Lauda.

²¹ Porto Editora. (n.d.) Legenda. In *Infopédia*. Retrieved February 2, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/legenda>

²² Porto Editora. (n.d.) Lenda. In *Infopédia*. Retrieved February 2, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/lenda>

3. Pele merupakan legenda hidup bagi Brasil. (lms.unhas.ac.id, collected on 05/02/2014)

Pele was Brazil's living legend.

Interestingly, more than *legenda* in the context of historical tale, in the corpus the word was much more often used to refer to someone legendary, especially in sport-related articles or news, especially in football. This was also proven by looking at the word graph provided on the entry page of Corpora Collection, as seen on Figure 19 below:



Figure 19 Word graph of the loanword *legenda* (legend) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=legenda (retrieved on January 21, 2022)

Even without knowing any Indonesian, anyone with a slight knowledge on football would realize that the word *legenda* often revolves around said sport, especially seeing the frequently-mentioned names Diego Maradona, together or separately, as well as Pele. It is also often mentioned together with Argentina, *sepakbola* (football) and *rekor* (record), all of which are very related to the football legend, Diego Maradona.

As far as we could see on the first hundred occurrences of the word, we searched without encountering with the use of *legenda* in the context of mapping or similar, consequently we would conclude that the word is rather infrequently used in that context and thus, it is correctly categorized in the semantic field of entertainment.

3.4.3. Warfare

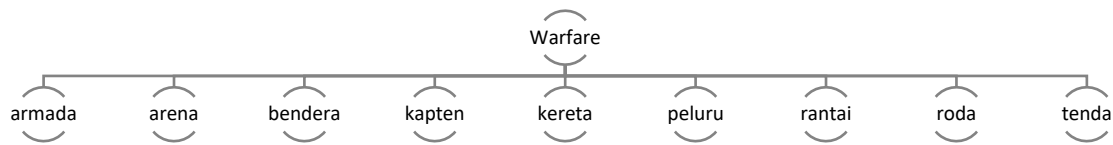


Figure 20 Semantic field of Portuguese loanwords in Indonesian: warfare

The next semantic field is warfare (see Figure 20), in which we included nine loanwords, namely *armada* (fleet), *arena* (arena/field), *bendera* (flag), *kapten* (captain), *kereta* (cart/train), *peluru* (bullet), *rantai* (chain), *roda* (wheel) and *tenda* (tent).

One word that caught our attention was *armada*. Initially, it meant a fleet of ships²³, be it in war or commerce. It is still used in that context in modern-day Indonesian as well, although it has gained an additional meaning of a group of something, not necessarily ships, according to KBBI²⁴. This we can prove with sample sentences from Corpora Collection:

1. TNI AD tampak tengah mengembangkan armada heli angkut pasukannya. (kuliah-karyawan.mercubuana.ac.id, collected on 31/01/2014)

It seems the Indonesian Army is developing its transportation helicopter fleet.

2. Dari situlah nama armada semut disematkan. (groups.yahoo.com, collected on 08/05/2012)

That is how it received the name of fleet of ants.

Tak hanya lesu, mental armada asuhan Walter Mazzarri itu pun juga terluka. (terkini.bbc.web.id, collected on 03/02/2014)

Not only tired, but Walter Mazzarri's squad is also mentally hurt.

As we can see, the meaning of *armada* became a fleet of other things such as helicopter, a squad of football players and even ants, surprisingly, although the latter seems to be metaphorical in meaning.

We also decided to check the word graph of the above-mentioned word to ensure the use of the word, and obtained the following (see Figure 21):

²³ Porto Editora. (n.d.) Legenda. In *Infopédia*. Retrieved January 4, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/armada>

²⁴ KBBI Daring. (n.d.) Armada. In *KBBI Daring*. Retrieved January 24, 2022, from <https://kbbi.kemdikbud.go.id/entri/armada>

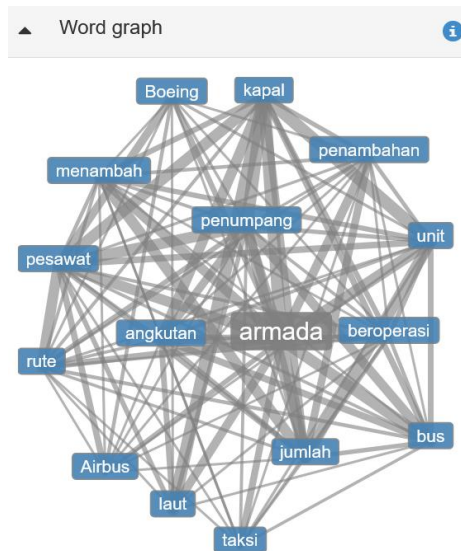


Figure 21 Word graph of the loanword *armada* (fleet of ships) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=armada (retrieved on January 24, 2022)

Based on the word graph above (see Figure 21), we can confirm that the word *armada* does not really relate with ships anymore. In fact, the only co-occurrence specifically related to ships is *laut* (sea), while the rest are about land or air transportation such as *pesawat* (airplane), *taksi* (taxi), *bus* (bus), as well as the airplane types *Boeing* and *Airbus*, or some more general words such as *rute* (route), *jumlah* (amount, quantity), or *penumpang* (passenger).

3.4.4. Religion

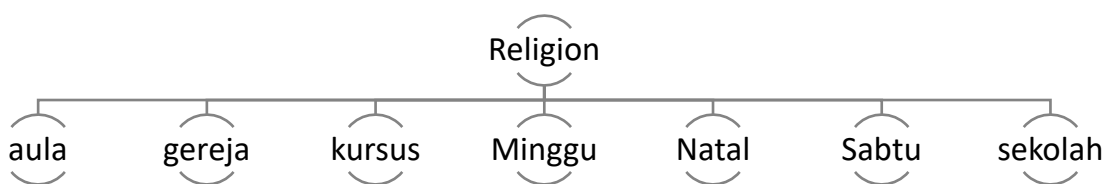


Figure 22 Semantic field of Portuguese loanwords in Indonesian: religion

The next semantic field we decided on was religion (see Figure 22), which included seven words: *aula* (class, classroom), *gereja* (church), *kursus* (course), *Minggu* (Sunday), *Natal* (Christmas), *Sabtu* (Saturday) and *sekolah* (school). We know that the Portuguese not only traded in Indonesia, but also

spread Catholicism by sending missionaries, especially in the Moluccas islands in the eastern part of the archipelago²⁵. That is why, these chosen words are relevant in the semantic field.

However, aside from *gereja* and *Natal*, nowadays the words above are not necessarily used in the field of religion anymore, but in a wider context. If initially *sekolah* was related to Sunday schools built and run by the missionaries, in modern-day Indonesian they simply mean school in general, which we can see in the context of the following sample sentences:

1. Tiba-tiba bel sekolah berbunyi. (andriandriant.blogspot.com, collected on 08/05/2012)
Suddenly, the school bell rang.
2. Pagi-pagi sekali aku berangkat sekolah. (www.lokerseni.web.id, collected on 31/01/2014)
I go to school very early in the morning.
3. Selama jam sekolah Ibu Ryan selalu menunggu Ryan disekolah. (disdik.jakarta.go.id, collected on 31/01/2014)
During school time, Ryan's mother always waits for him at school.

As we can see, all samples above speak of school in general terms. Of course, the religious term *sekolah Minggu* (Sunday school) still exists, but it is not notably so, especially considering Indonesia is a country where the 86,7% of the population are Muslim²⁶ (despite not being an Islamic country ("Islam in Indonesia", 2022)²⁷).

However, we were in a doubt whether *kursus* (course) really was used in religious terms initially, since nowadays it is also used in a wider, more general terms. It usually means materials or skills taught or institutions teaching those skills outside school time. This description we can see from the following sample sentences in the corpus:

1. Saya berminat kursus bahasa jerman. (pend-bhsindonesias2.umm.ac.id, collected on 01/02/2014)
I am interested in the German language course.
2. Jika diperlukan, ambilah kursus singkat yang tidak terlalu menyita waktu, bila Anda juga seorang pekerja kantoran. (forum.prodigy.co.id, collected on 07/02/2014)
If needed, take a short course that does not require much of your time, if you also work.

²⁵ One notable missionary in Moluccas was Francis Xavier (PT: Francisco Xavier), who was able to catholicize the initially Muslim population in Ambon. He did so by staking big crosses in villages, followed by baptizing the people in those villages. He also taught the youngsters in catechesis and built numerous Catholic schools around the Portuguese fort. (Wijaya D.N., Wahyudi, D.Y., Umaroh, S.Z., 2021, p. 79-81)

²⁶ Indonesia.Go.Id. Portal Informasi Indonesia. (n.d). Agama. <https://www.indonesia.go.id/profil/agama> (retrieved on January 24, 2022)

²⁷ Islam in Indonesia (2022, January 24). In *Wikipedia*.
https://en.wikipedia.org/w/index.php?title=Islam_in_Indonesia&oldid=1066621470

3. Video ini adalah dokumentasi kursus dan pelatihan yang dilaksanakan oleh LKP Bintang Terang. (website.dinar.co.id, collected on 08/02/2014)

This video is a documentation of the course and training held by LKP Bintang Terang.

Considering the samples above, we see that it is indeed used in general terms. Therefore, at this point we can only hypothesize about it having religious origins and that it was originally tied to school, and can probably figure out more of its etymology with further studies. Unfortunately, the World Loanword Database also does not have this word included in the sample, so it is not possible to cross-check its etymology with them.

While analyzing the words, we also became uncertain of the word *aula* and whether it was really brought into Indonesian in a religious context, especially because modern-day Portuguese has the same word with a different meaning. In the Portuguese monolingual dictionary Infopédia, the word *aula* means:

Aula²⁸ (n, f)

1. *recinto onde se recebe uma lição*
2. *classe*
3. *lição*
4. *palácio de príncipe*
5. *espécie de conselho que o rei convocava quando entendia, nos primeiros tempos da monarquia portuguesa*

The way the word is defined, it means a class, a classroom, a prince's palace, or some kind of councilor in the early Portuguese monarchy, although the most common use is class as seen on definition 2. Meanwhile, in Indonesian, it means a big, multifunctional room, like an auditorium. Our hypothesis would be that the word used to be used as class or classroom in the schools the missionaries built in Indonesia, hence we decided to include it in this semantic field.

²⁸ Porto Editora. Aula. In *Infopédia*. Retrieved January 25, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/aula>

3.4.5. Household

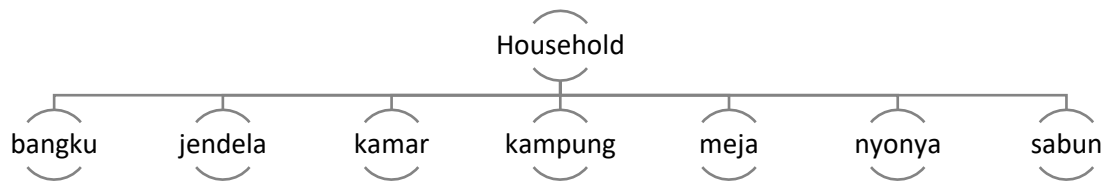


Figure 23 Semantic field of Portuguese loanwords in Indonesian: household

Next, our chosen semantic field is household (see Figure 23), which includes seven Portuguese loanwords, namely: *bangku* (bank to sit on), *jendela* (window), *kamar* (room), *kampung* (village), *meja* (table), *nyonya* (missus, married lady) and *sabun* (soap).

We decided to take a look at the word graph of the word *kampung* on Corpora Collection and obtained the following relations (see Figure 24):

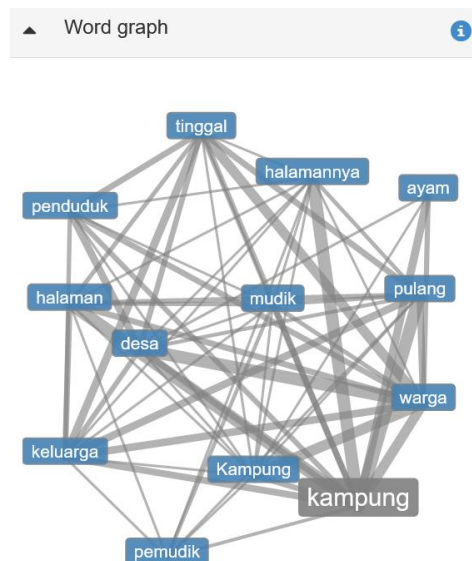


Figure 24 Word graph of the loanword *kampung* (village) on Corpora Collection, from https://corpora.uni-leipzig.de/en/res?corpusId=ind_mixed_2013&word=kampung (retrieved January 25, 2022)

On the word graph above (see Figure 24) we see that the word *kampung* often co-occurs with *warga* (citizens), *keluarga* (family), *mudik* (to return to the hometown) and *pemudik* (people returning to hometown). There is also *pulang* (to return home) that makes the compound *pulang kampung*

(synonym to *mudik*) and *halaman* that makes the compound *kampung halaman* (hometown). However, we still included the word into this semantic field because the co-occurrences imply one.

3.4.6. Clothing

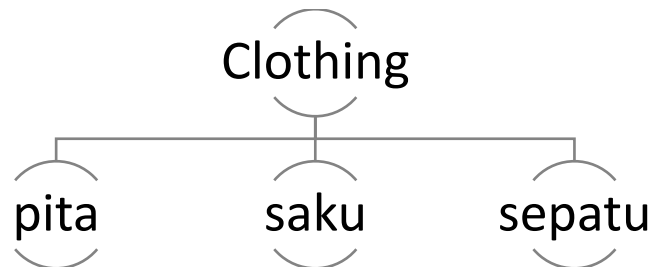


Figure 25 Semantic field of Portuguese loanwords in Indonesian: clothing

The last semantic field in our analysis is clothing (see Figure 25), with three words: *pita* (ribbon), *saku* (pocket) and *sepatu* (shoe).

Currently, the words in Indonesian have not changed in meaning from the donor words in Portuguese, with the exception of *saku*. In Portuguese, the donor word *saco* means bag, while in Indonesian, *saku* means pocket, so the meaning has altered during the adaptation.

We would like to emphasize this hypothesis using some sample sentences from Corpora Collection:

1. Kamera saku terbaru yang diperkenalkan Sony Indonesia bisa dikatakan tertipis saat ini. (en.uin-malang.ac.id, collected on 03/02/2014)
The newest pocket camera Sony Indonesia introduced could be considered the thinnest camera to date.
2. Tadi aku menyimpannya di saku bajuku ini, tapi sekarang tak ada. (kisahbebe.blogspot.com, collected on 08/05/2012)
I kept it in my shirt's pocket earlier, but now it's gone.
3. Ismi mengaku tidak mengembalikan uang tersebut karena dianggap sebagai uang saku. (waspada.co.id, collected on 04/02/2014)
Ismi admitted not having returned the money, because she considered it allowance.

From these samples, we can see that the words really literally mean pocket. However, we also have the expression *uang saku* as seen in example 3, which means allowance. This expression is very commonly used, in fact many sentences in the corpora use the word in this context.

From the semantic fields we decided on above, we could conclude that every single word among the forty-seven most frequent Portuguese loanwords can belong to one, despite some of them having left us with a moment of doubt whether they really fit where we had assigned them.

As seen on the explanation of each semantic field, we had also decided to assign the semantic fields diachronically, as we mentioned at the beginning of this sub-chapter, that we wanted to see how these loanwords were adapted into the then-Malay. We think it is important to know the meanings during the adaptation process, because through centuries, those meanings may have broadened, narrowed, or even changed.

We could see that example in chapter 3.4.6. in the case of the loanword *saku* (pocket), loaned from the word *saco*, which means bag in general, however in modern Indonesian, its specifically means pocket and not just any bag; which means that the meaning has narrowed in comparison to the loaned word.

As for how or why these semantic changes may have occurred in Indonesian, they would require further studies. However, we could say that there are motivations as to why the changes occurred and mechanisms that study how the changes occurred (Traugott, 2017, p. 8).

To conclude our analysis on semantic fields and the loanwords in our study in general, we put them together on one graph for easier viewing, as seen below (see Figure 26):

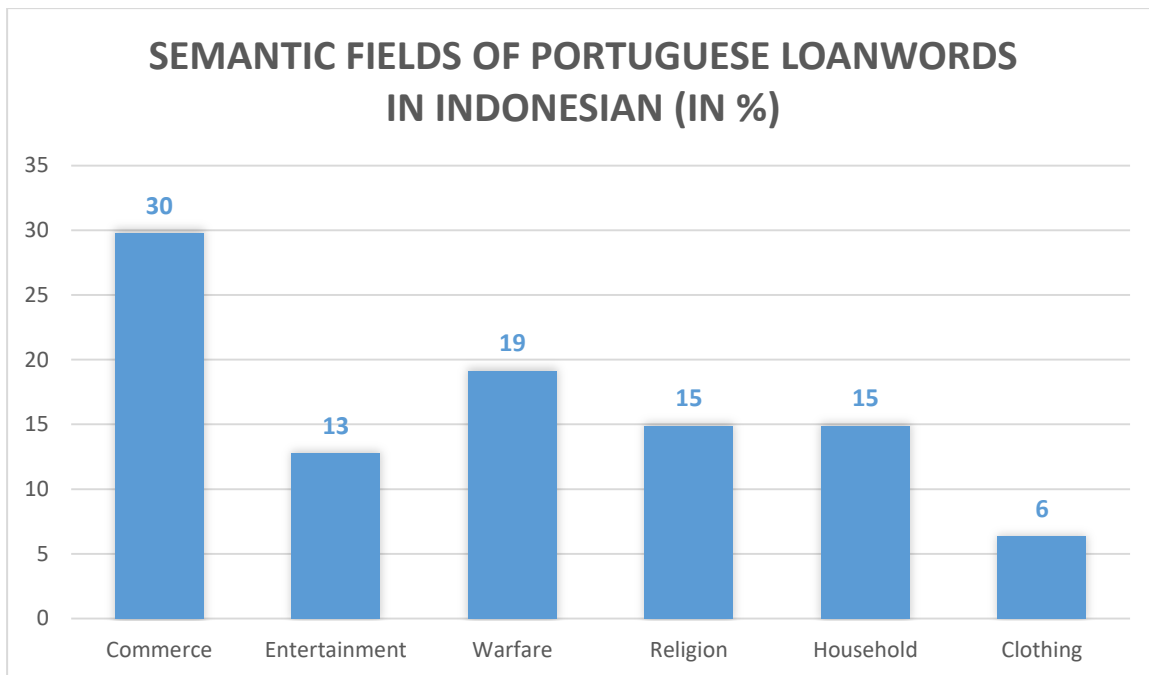


Figure 26 Semantic Fields of Portuguese Loanwords

To summarize, of the total forty-seven words, the semantic field commerce has fourteen words, comprising 29.79% of the words, entertainment has six words or 12.77%, warfare has nine or 19.15%, religion and household have seven words or 14.9% respectively, and last but not least, clothing has the remaining three words, making 6.38% of the total percentage.

Looking back to Chapter 1 on Tadmor's research on Indonesian loanwords, he had identified that these four semantic fields have a high loanword content: Religion and belief (70% borrowed vocabulary), the Modern world (66,4%), Clothing and grooming (55,6%) and Law (51,4%) (Haspelmath, M. & Tadmor, U., 2009, p. 700). We discovered that our analysis on Portuguese loanwords specifically did not deviate from those four semantic fields, considering that in our case the largest ones include commerce, warfare, household and religion.

This is a logical result, considering those were the major activities the Portuguese conducted while in the archipelago. They traded goods, fought for spices and land with both locals and merchants of other nationalities, and also spread their religion. That was how words related to those field were frequently used, then left behind, adapted into the local language and continue being used until today.

After this analysis, we have also gathered that most of the loanwords are nouns. This result is in line with what Peter Eisenberg said about the hierarchy of loanwords, in which "Noun > Adjective > Verb >

Adverb > Preposition > other function words > Derivation > Inflection > Sound.” (2018, p. 28) This means that nouns are easiest to be loaned after a language contact than other parts of speech.

Chapter 4: Dictionary Representation

Now that we have compiled a list of known Portuguese loanwords using the resources in Chapter 2 and analyzed them for frequency, connotation and semantic field in Chapter 3, we could use that list in order to try making a representation on how they would look in a dictionary, because ultimately, the long-term goal would be to create a dictionary of Portuguese loanwords in Indonesian.

4.1. Early Representation Idea

Back when the assignment on Dictionary Planning and Production was written in 2020, we included an example on how a lemma would look like in a dictionary of loanwords as part of the user's guidelines. Coincidentally, one of the examples was made for a Portuguese loanword, as seen on Figure 27 below:

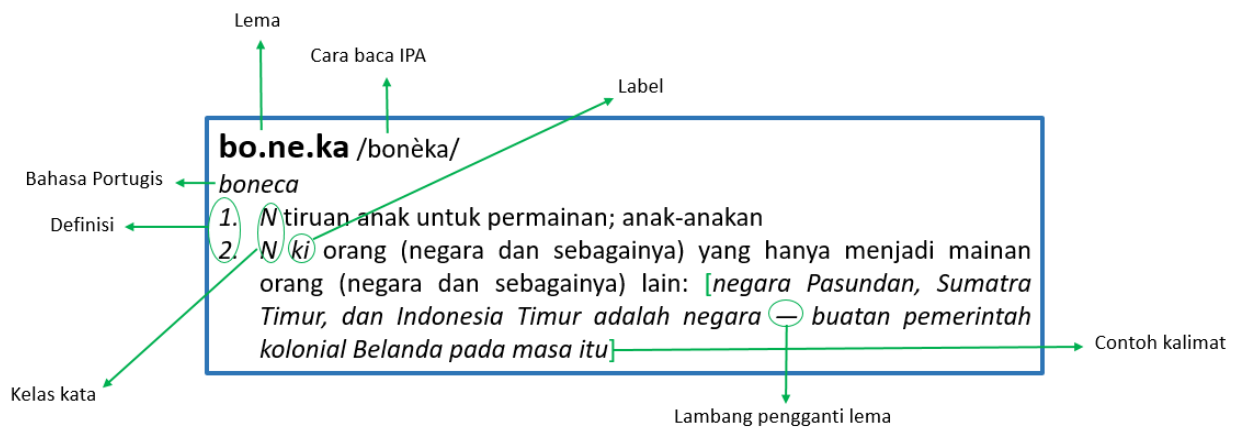


Figure 27 Example of dictionary entry for a Portuguese loanword (Yudo, 2020)

The example on that previous paper as seen on Figure 27 used the word *boneca* (doll). The idea was to base our dictionary entries on KBBI Daring's entry (see Figure 2 in Chapter 2), with the following information:

- First row: lemma written in bigger font and bold, followed by IPA reading.
- Second row: the source word for the lemma.
- Third row: definition, including sense(s), part of speech, label if any and example sentence.

In this early representation idea, the target audience of the dictionary was students, both Indonesians studying philology or literature especially, as well as foreigner studying Indonesian with at least intermediate level of language knowledge. Furthermore, "laypeople who have interest in Indonesian

linguistics may also enjoy this dictionary series to broaden their repertoire.” (Yudo, 2020) With that said, the dictionary series is meant not only for studying, but also for reading purposes, therefore, it was thought to be designed as simple as possible.

4.2. Representation Idea after Analysis

While we think the initial idea of microstructure was good enough for the audience it was targeted at, after the study we did this time, we determined that the microstructure in a dictionary of loanwords should include more information based on our analysis, namely naming the semantic field and frequency. With the semantic field, we can tell the users in which way or field the words were introduced into our language, while the frequency determines how easily a word can be identified and remembered by a language user (Divjak, D. & Caldwell-Harris, C., 2015, p. 56). In the case of our study, the frequency of Portuguese loanwords can also determine their level of commonness within the Indonesian vocabulary.

The semantic field will be named according to the analysis we did in Chapter 3.4. However, since we only analyzed the 47 most frequent words (see Table 3) and yet there were 126 words that we have compiled into a list of known Portuguese loanwords in Indonesian (see Table 2), the rest of the list will have to be categorized into their semantic fields accordingly.

As for the frequency, it does not need to be written in their specific ranks, however, since the numbers are not absolute, considering the possible deviations from typos, suffixes and homonymy, as mentioned on Chapter 2 and Chapter 3. Instead, we would categorize the words into different degrees of frequency class, as follow:

- Words ranking 1,000 and above: highly frequent (ID: *sangat sering*)
- Words ranking 10,000 and above: frequent (ID: *sering*)
- Words ranking 100,000 and above: often (ID: *lazim*)
- Words ranking 1,000,000 and above: seldom (ID: *jarang*)
- Words ranking less than 1,000,000: very rare (ID: *sangat jarang*)

We decided to choose six representative words based on the semantic fields we worked on in Chapter 3, of the words with frequency rank 10,000 and above. We came upon this latter decision because only those forty-seven words have been analyzed and we determined that choosing something from outside that list would require more study. Hence, we chose one word from each semantic field, as follow:

- Semantic field commerce: *palsu* (false, fake)
- Semantic field entertainment: *boneka* (doll)
- Semantic field warfare: *armada* (fleet)
- Semantic field religion: *sekolah* (school)
- Semantic field household: *meja* (table)
- Semantic field clothing: *saku* (pocket)

Since the dictionary will be in Indonesian, we also translated the semantic fields into: *perdagangan* (commerce), *hiburan* (entertainment), *peperangan* (warfare), *agama* (religion), *rumah tangga* (household) and *pakaian* (clothing).

Moreover, a dictionary of loanwords would normally include some etymological information or a history of the lemma and that information we will incorporate as well, based on the Portuguese (and Spanish) etymological dictionaries we described in Chapter 2. Other than that, information such as definition(s) and part of speech will still be taken from KBBI Daring, while the example in sentence will be taken from Corpora Collection.

The dictionaries that will be used as we have mentioned in Chapter 2 are namely the *Diccionario da lingua portugueza composto pelo padre D. Rafael Bluteau by António de Morais Silva (1789)*, the *Dicionário Etimológico da Língua Portuguesa by José Pedro Machado (1956)*, the *Diccionario crítico etimológico castellano e hispánico by Joan Corominas (1954)* and *Dicionário Houaiss da língua portuguesa by António Houaiss (2003)*. The first dictionary serves the purpose of collecting the definition, while the other three are where we will collect the etymological information from. And the idea is to put the information found in each dictionary side-by-side in a table for an easy readability.

Aside from the additional information described above, we will also modify the layout of the microstructure, so that despite having plenty of information, it will still be understandable and readable for the same target audience (mainly students), but also for laypeople interested in the field.

The idea is that in the actual dictionary, the preface will also include a user's guideline, in which we will show images of sample dictionary entries of the dictionaries we used as reference, in order for users to know where we take the references from, especially since the planned dictionary will be in Indonesian, which means that the information we cite will be translated.

All the above considered, below we will show our proposed dictionary entries one-by-one. But beforehand, we will show the entry template translated into English for the non-Indonesian-speaking in Figure 28 seen below.

[<i>lemma, with syllable-separating dots and written in bold</i>]			
Definition: [<i>definition(s), accompanied by part of speech and label if any</i>]			
Example in sentence: [<i>taken from Corpora Collection, written in italics and the lemma underlined</i>]			
Source word: [<i>written in italics for being a foreign word</i>]			
Semantic field:			
Frequency class:			
Etymology:			
Morais	Machado	Corominas	Houaiss

Figure 28 Template of the microstructure of our dictionary of Portuguese loanwords

Despite the translated template in Figure 28, in the samples of dictionary entries that we will show below, we will also include the translation of each lemma for easier understanding of this representation.

4.2.1. Palsu

pal.su

Definisi:

1. a Tidak tulen; tidak sah; lancing (tentang ijazah, surat keterangan, uang, dan sebagainya)
2. a Tiruan (tentang gigi, kunci, dan sebagainya)
3. a Gadungan (tentang polisi, tentara, wartawan, dan sebagainya)
4. a Curang, tidak jujur (tentang permainan dan sebagainya)
5. a Sumbang (tentang suara dan sebagainya)

Contoh kalimat: *Pencuri diduga menggunakan kunci palsu.*

Kata asal: *falso*

Bidang semantik: perdagangan

Tingkat frekuensi: sering

Etimologi:

Morais	Machado	Corominas	Houaiss
Kebalikan dari benar, tidak benar; tiruan; palsu.	Dari bahasa Latin <i>falsu-</i> [palsu, tiruan, penipu, gadungan, pembohong]	Dari bahasa Latin <i>falsus</i> , terutama bentuk pasif dari <i>fallĕre</i> 'menipu'. Dok. 1, paruh kedua abad X, Glosas Silenses.	Dari bahasa Latin <i>falsus, a, um</i> [salah, tidak benar], bentuk pasif dari <i>falsō, is, fefellī, falsum, ěre</i> [salah].

English Translation:

pal.su

Definition:

1. a Not authentic; not valid (certificate, letter, money, etc.)
2. a Imitation or duplicate (teeth, keys, etc.)
3. a Impostor, copycat (police, military, journalist, etc.)
4. a Unfair, dishonest (game, etc.)
5. a discordant, off-key (voice, etc.)

Example in sentence: *The robber is suspected to have used a duplicate key.*

Source word: *falso*

Semantic field: commerce

Frequency class: frequent

Etymology:

Morais	Machado	Corominas	Houaiss
Opposite of correct, not correct; falsified; fake.	From Latin <i>falsu-</i> [not true, imitation, liar, impostor]	From Latin <i>falsus</i> , esp. passive form of <i>fallĕre</i> 'to lie'. Doc. 1, 2 nd half, 10 th century, Glosas Silenses.	From Latin <i>falsus, a, um</i> [wrong, not real], passive form of <i>falsō, is, fefellī, falsum, ěre</i> [to be wrong].

4.2.2. Boneka

bo.ne.ka

Definisi:

1. n Tiruan anak untuk permainan; anak-anakan.
2. n ki Orang (negara dan sebagainya) yang hanya menjadi mainan orang (negara dan sebagainya) lain.

Contoh kalimat: *Sang boneka kodok selalu menemani tidur malamnya.*

Kata asal: *boneca*

Bidang semantik: hiburan

Tingkat frekuensi: sering

Etimologi:

Morais	Machado	Corominas	Houaiss
Figur yang meniru perempuan, terbuat dari kertas, kain, dll.	Dari bahasa Spanyol <i>muñeca</i> dan bahasa Galicia <i>moneca</i> ., tampaknya bahasa-bahasa Iberia terhubung dengan kata <i>ninna</i> dan <i>nonna</i> , kata panggilan anak-anak Romawi untuk orang yang disayang. Ada catatan yang menyebutkan <i>monna</i> , panggilan sayang seorang suami pada istrinya.	[benda berbentuk wanita yang digunakan sebagai mainan]. Berasal dari zaman pre-Romawi, bersaudara dengan bahasa Portugis <i>boneca</i> , yang pada awalnya bernama <i>moñeca</i> , sama-sama berawal dari kata <i>bonnīcca</i> , yang mungkin berasimilasi dari <i>bodinicca</i> , turunan dari bahasa Kelt <i>bodīna</i> .	Dari bahasa Latin <i>ninna</i> dan <i>nonna</i> , panggilan sayang yang digunakan anak-anak Romawi untuk memanggil orang yang disayang; catatan abad ke-14 <i>moneca</i> , 1562 <i>boneca</i> .

English Translation:

bo.ne.ka

Definition:

1. n Child imitation as toy; child-like.
2. n ki A person (or country, etc.) used as puppet of others (person, country, etc.).

Example in sentence: *The frog doll always accompanies her to sleep.*

Source word: *boneca*

Semantic field: entertainment

Frequency class: frequent

Etymology:

Morais	Machado	Corominas	Houaiss
A figure resembling a woman, made of paper, fabric, etc.	From Spanish <i>muñeca</i> and Galician <i>moneca</i> , it seems the Iberian languages are connected to the word <i>ninna</i> and <i>nonna</i> , a diminutive used by Roman children to call their loved ones. A note mentions <i>monna</i> , an endearing term used by a husband to his wife.	[a woman-looking figure used as toy]. From the pre-Roman times, related to Portuguese word <i>boneca</i> , which at the beginning was called <i>moñeca</i> , both originally from the word <i>bonnīcca</i> , maybe assimilated from <i>bodinicca</i> , a derivative of Celtic word <i>bodīna</i> .	From Latin <i>ninna</i> and <i>nonna</i> , a diminutive used by Roman children to call their loved ones; note from 14 th century <i>moneca</i> , 1562 <i>boneca</i> .

4.2.3. Armada

ar.ma.da

Definisi:

1. n rombongan (pasukan) kapal perang
2. n Rombongan kapal dagang
3. n Rombongan suatu kesatuan

Contoh kalimat: *Waktu itu Kadiri memiliki armada laut yang cukup tangguh.*

Kata asal: *armada*

Bidang semantik: peperangan

Tingkat frekuensi: sering

Etimologi:

Morais	Machado	Corominas	Houaiss
Sejumlah kapal untuk perang. [<i>Andar d'armada e alguma paragem</i>] Berlayar, menepi, berhenti, untuk menanti atau mengamati musuh, menjaga pantai, atau aksi militer laut lainnya.	[<i>Armadilha de caça</i>], dari kata <i>armar</i> . Dalam arti [sekumpulan kapal], dari bahasa Italia <i>armata</i> (abad XIII), dengan pengaruh dari adjektiva <i>armado</i> . Pengaruh dari bahasa Spanyol <i>armadilla</i> , <i>armandija</i> .	<i>Armada</i> [1476, Woodbr.; APal. 297b]. <i>Armandijo</i> [A. de Cartagena, †1456]	Dari kata <i>armado</i> ; menurut catatan tahun 1130 untuk istilah [<i>armadilha de caça</i>] dan 1317 untuk [sekelompok kapal].

English Translation:

ar.ma.da

Definition:

1. n A fleet of war ships

2. n A fleet of commerce ships
3. n A group of something

Example in sentence: *At the time, Kadiri had quite a strong armada.*

Source word: *armada*

Semantic field: warfare

Frequency class: frequent

Etymology:

Morais	Machado	Corominas	Houaiss
A fleet of ships for war. [<i>Andar d'armada e alguma paragem</i>] To sail, to come close to coast, to stop in order to wait or observe the enemy, to guard the coast, or other maritime military act.	[<i>Armadilha de caça</i>], from the word <i>armar</i> . In the sense of [a fleet of ships], from Italian <i>armata</i> (13th century), influenced by the adjective <i>armado</i> . Influence from Spanish <i>armadillo</i> , <i>armandija</i> .	<i>Armada</i> [1476, Woodbr.; APal. 297b]. <i>Armandijo</i> [A. de Cartagena, †1456]	From the word <i>armado</i> ; according to a note from 1130 for the expression [<i>armadilha de caça</i>] and 1317 for [a fleet of ships].

4.2.4. Sekolah

se.ko.lah

Definisi:

1. N Bangunan atau lembaga untuk belajar dan mengajar serta tempat menerima dan memberi pelajaran (menurut tingkatannya).
2. N Waktu untuk pertemuan ketika murid diberi pelajaran.
3. N Usaha menuntut kepandaian (ilmu pengetahuan); pelajaran; pengajaran.

Contoh kalimat: *Pagi-pagi sekali aku berangkat sekolah.*

Kata asal: *escola*

Bidang semantik: agama

Tingkat frekuensi: sangat sering

Etimologi:

Morais	Machado	Corominas	Houaiss
Rumah tempat diajarkannya membaca, menulis, menari, membela diri.	Dari bahasa Yunani <i>sholē</i> [pemberhentian, istirahat; pekerjaan tentang belajar, pekerjaan bijak, pendidikan; berhubungan dengan budaya; tempat belajar, sekolah; hasil dari pelajaran], dari bahasa Latin <i>schola</i> [kegiatan yang didedikasikan pada pelajaran, kelas, konferensi; tempat mengajar, sekolah; galeri; ruang tunggu di kamar mandi; perusahaan]	Dari Bahasa Latin <i>schōla</i> [pelajaran, sekolah] dan Bahasa Yunani <i>σχολή</i> [waktu kosong, pelajaran, sekolah]. Dokumen pertama mozárabe tahun 1192 (Oelschl.); <i>escuela</i> , Berceo.	Dari Bahasa Latin <i>schōla, ae</i> [tempat di kamar mandi untuk menunggu giliran; pekerjaan kesusastaaran, pelajaran; sekolah, kelas] atau dari Bahasa Yunani <i>skholē, és</i> [istirahat, waktu luang; pelajaran; pekerjaan bagi orang yang memiliki waktu kosong, bebas dari tugas sipil; tempat belajar].

English Translation:

se.ko.lah

Definition:

1. n Building or institution for teaching and studying, as well as receiving and giving lessons (according to their level).
2. n Meeting time when students are given lessons.
3. n The effort of looking for knowledge; lesson; teaching.

Example in sentence: *I go to school very early in the morning.*

Source word: *escola*

Semantic field: religion

Frequency class: highly frequent

Etymology:

Morais	Machado	Corominas	Houaiss
A house where reading, writing, dancing and self-defense is taught.	From Greek <i>sholē</i> [stop, to rest; occupation related studying, a wise occupation, knowledge; related to culture; a place to study, school; result of studying], from Latin <i>schola</i> [activity dedicated to studying, class,	From Latin <i>schōla</i> [lesson, school] and Greek <i>σχολή</i> [free time, lesson, school]. First doc. 1192 (Oelschl.); <i>escuela</i> , Berceo.	From Latin <i>schōla, ae</i> [waiting room in baths; literary occupation, lesson, school, class] or Greek <i>skholē, és</i> [to rest, free time; lesson; job for people with free time and no need to do civil work; a place to

	conference; a teaching place, school; gallery; waiting room in baths; company]		study].
--	--	--	---------

4.2.5. Meja

me.ja

Definisi:

n Perkakas (perabot) rumah yang mempunyai bidang datar sebagai daun mejanya dan berkaki sebagai penyangganya (bermacam-macam bentuk dan gunanya).

Contoh kalimat: *Guru mendatangi meja siswa satu-persatu.*

Kata asal: *mesa*

Bidang semantik: rumah tangga

Tingkat frekuensi: sering

Etimologi:

Morais	Machado	Corominas	Houaiss
Perabot untuk pelayanan di rumah, tempat meletakkan makanan ketika makan malam.	Dari bahasa Latin populer * <i>mēsa</i> , Latin klasik <i>mensa</i> [meja. Kiasan: makanan, hidangan; undangan; tamu; balkon, meja bank; meja tempat meletakkan obyek sacral atau meja persembahan; panggung tempat para budak dijual].	Dari bahasa Latin <i>mensa</i> id. Dokumen pertama: origenes del idioma (doc. 978 Oelschl, etc).	Dari bahasa Latin <i>mensa, ae</i> [meja untuk makan, meja kuil tempat meletakkan sesajen dan alat-alat lainnya, meja makan,]. 1395 <i>messsa</i> , abad ke-14 <i>mēsa</i> , 1716 <i>mensa</i> .

English Translation:

me.ja

Definition:

n House furniture that has a flat surface and legs as supporters (it can have different shapes and purposes).

Example in sentence: *The teacher goes to the students' desks one by one.*

Source word: *mesa*

Semantic field: household

Frequency class: frequent

Etymology:

Morais	Machado	Corominas	Houaiss
Furniture used for service at home, a place to put food on during supper.	From Latin * <i>mēsa</i> , Classical Latin <i>mensa</i> [table. Figurative: food, meal; invitation; guest; balcony, teller; table to put sacred or offering on; stage where slaves are sold].	From Latin <i>mensa</i> id. First doc.: <i>origenes del idioma</i> (doc. 978 Oelschl, etc).	From Latin <i>mensa, ae</i> [table for having meals, temple table to put offerings and other attributes on, dining table]. 1395 <i>messas</i> , 14 th century <i>mēsa</i> , 1716 <i>mensa</i> .

4.2.6. Saku

sa.ku

Definisi:

n Kantong (pada baju, celana, rok, dan sebagainya).

Contoh kalimat: *Bagian dalam tas terbuat dari kain coklat tua dan ada sebuah saku dalam tanpa penutup retsleting.*

Kata asal: *saco*

Bidang semantik: pakaian

Tingkat frekuensi: sering

Etimologi:

Morais	Machado	Corominas	Houaiss
Tempat penyimpanan yang terbuat dari kain atau kulit, terdiri dari dua potongan persegi empat yang dijahit pada tiga sisi; satu sisi dibiarkan terbuka sebagai mulut untuk memasukkan barang untuk dibawa.	Dari bahasa Latin <i>saccu-</i> [kantong, kain untuk menyaring; dalam bahasa Latin eklestik, pakaian kasar yang terbuat dari kulit kuda (C.G.L., III, 197, 10)].	Dari Bahasa Latin <i>saccus</i> [kantong terigu, uang, dll.] [pakaian kasar], Bahasa Yunani <i>σάχος</i> id. dan [pembunngkus], kata dengan asal Fenisia. Dok. pertama: Berceo.	Dari bahasa Latin <i>saccus, i</i> [kantong untuk menyaring], dari bahasa Yunani <i>sákkos, ou</i> [kain kasar], kata serapan dari Bahasa Ibrani <i>sak</i> [kain kasar untuk berbagai fungsi (alas, penutup)]

English Translation:

sa.ku

Definition:

n Pocket (in shirt, pants, skirt, etc.)

Example in sentence: *The inner part of the bag is made of a dark brown fabric and has an inner pocket without zipper.*

Source word: *saco*

Semantic field: clothing

Frequency class: frequent

Etymology:

Morais	Machado	Corominas	Houaiss
A safekeeping bag made of fabric or leather, which consists of two square pieces and sewn on its three sides; one side is left open as opening to insert things it will carry.	From Latin <i>saccu-</i> [pocket, filtering fabric; in Ecclectic Latin: rough cloth made of horse skin (C.G.L., III, 197, 10)]	From Latin <i>saccus</i> [bag of flour, money, etc.] [rough cloth], from Greek <i>σάχος</i> id. [wrapping], originally from Fenician. First doc.: Berceo.	From Latin <i>saccus, i</i> [filtering bag], from Greek <i>sákkos, ou</i> [rough cloth], loaned from Hebrew <i>sak</i> [rough cloth for multiple purposes (floor covering, cover)]

Conclusion

As mentioned on Chapter 0.2., we based our research on the following research questions:

1. Identify the Portuguese loanwords frequently found within the Indonesian language.
2. Why are most of the high-frequency Portuguese loanwords nouns?
3. Do Portuguese loanwords in Indonesian language carry any connotation?
4. Which semantic fields do the high-frequency Portuguese loanwords belong to?

Thus, in our research, we firstly discovered that Portuguese loanwords do not make the 50 most frequent words in Indonesian, which is a list that consists of words that are mostly conjunctions and a few pronouns. We discovered that 16% of them are known to be loanwords, but they were either of Sanskrit or Arabic origin.

Then, we looked at the loanwords from a different point of view. We compiled known Portuguese loanwords from various available glossaries and obtained a list of 126 words. We checked for their frequencies one by one, taking into account that some of the words are homonyms, so that the frequencies may deviate. Next, for easier analysis, out of those 126 words we only picked words of frequency rank 10,000 and higher, obtaining a smaller list of 47 words, of which eight (17.07%) rank 1,000 and above, twenty-four (51.06%) rank 5,000 and above and fifteen (31.91%) rank 10,000 and above. It was also worth noting that most of the Portuguese loanwords are nouns, a fact that goes in line with the loanword hierarchy, in which nouns are sitting on top of loanwords.

We checked for any connotations of the 47 loanwords and based on KBBI team's explanation, words that are included in the dictionary generally don't have negative connotation. However, we analyzed our words and found out that while no particular negative connotation was found, in some cases it's the compounds that surge from those words that may have one. For example, the case of the word *palsu* (false, fake), the word itself can be found in KBBI Daring because it is not considered to have a negative connotation, and yet once it co-occurs with other words, such as *nabi* (prophet), making it *nabi palsu* (false prophet), it can cause a few raised eyebrows.

Subsequently, we analyzed the semantic fields of our list of 47 words. We decided on a few logical semantic fields that could categorize those words, namely commerce, entertainment, warfare, religion, household and clothing. We found that the semantic field of commerce has fourteen words (29.79%), entertainment has six words (12.77%), warfare has nine words (19.15%), both religion and household

have seven words (14.9%) and clothing has three words (6.3%). This somewhat coincides with the analysis on Indonesian in the World Loanword Database that the loanwords in Indonesian in general belong to the semantic fields religion and belief, modern world, clothing and grooming, and law.

As well, with the additional etymological information and how we proposed the loanwords to be represented in a dictionary of loanwords, this may become a starting point on an even further study on the history of the words. In this study, we took the information from the Portuguese side, but a future study may open new knowledge on the Indonesian side, which currently has not been done, as far as we know. The information gathered on this study may even be very useful for the time the KBBI team decides on picking up the Portuguese loanwords for their future updates, considering they are still ways to go from completing their etymological information.

The information may also be useful for further studies in the linguistic aspect. It would be very interesting to discover how the phonetics changed upon the adaptation of those loanwords in Indonesian. Based on the phonetical changes, it would even be possible to discover even more Portuguese loanwords among the Indonesian vocabulary. This we think possible, considering the long stay of the Portuguese in the archipelago, so it would be plausible for the number of loanwords to be higher.

Bibliography

- Abdurachman, P. R. (2008). *Bunga Angin Portugis di Nusantara*. Yayasan Pustaka Obor Indonesia
- Adam, A. (2020). *The Sejarah Melayu Revisited: A Collection of Six Essays*. Strategic Information and Research Development Centre.
- Buchi, E. (2015). Etymological dictionaries. In P. Durkin (Ed.), *The Oxford Handbook of Lexicography* (pp. 338 - 349). Oxford University Press.
- Divjak, D. & Caldwell-Harris, C. (2015). Chapter 3: Frequency and entrenchment. In E. Dabrowska & D. Divjak (Ed.), *Handbook of Cognitive Linguistics* (pp. 53-75). De Gruyter Mouton.
- Durkin, P. (2014). *Borrowed Words: A History of Loanwords in English*. Oxford University Press.
- Eisenberg, P. (2018). *Das Fremdwort im Deutschen*. De Gruyter Studium.
- Errington, J. J. (1986). Continuity and Change in Indonesian Language Development. *The Journal of Asian Studies*, 45(2), 329-353.
- Haspelmath, M. & Tadmor, U. (2009). *Loanwords in the World's Languages*. De Gruyter Mouton.
- Iriarte, A. (2001). *A Unidade Lexicográfica. Palavras, Colocações, Frasemas, Pragmatemas*. Centro de Estudos Humanísticos-Universidade do Minho.
- Jones, R. (general ed.), Grijns, C. D., De Vries, J. W. (Eds.) (2007). *Loan-words in Indonesian and Malay*. Compiled by the Indonesian Etymological Project. KITLV Press.
- Jumariam, M. T., Qodratillah, & Ruddyanto, C. (Eds.) (1996). *Senarai Kata Serapan dalam Bahasa Indonesia*. Departemen Pendidikan dan Kebudayaan.
- Pinto da França, A. & Rogers, R. (2014) *A influência portuguesa na Indonésia: Portuguese influence in Indonesia*. Ed. bilingue em português e inglês. ALIAC-Associação Luso-Indonésia para a Amizade e Cooperação, D.L.
- Traugott, E. Semantic Change. *Oxford Research Encyclopedia of Linguistics*. Retrieved 5 Jul. 2022, from <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-323>.

Van Coetsem, F. (1988) *Loan Phonology and the Two Transfer Types in Language Contact*. Foris Publications Holland.

Wijaya, D. N., Wahyudi, D. Y., & Umaroh, S. Z. (2021). *Nusa Ambon Silang Budaya: Pertemuan Jaringan Global Abad XVI-XVII*. Departemen Pendidikan dan Kebudayaan.

Yudo, Y. S. (2020). *Module 9: Dictionary Planning and Production* [Class Assignment]. European Master in Lexicography, University of Minho [unpublished].

Dictionaries

Corominas, J. & Pascoal, J. A. (1980-1983). *Diccionario crítico etimológico castellano e hispánico*. Gredos.

Houaiss, A. (2003). *Dicionário Houaiss da língua portuguesa*. Temas e Debates.

Machado, J. P. (1967). *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência.

Morais, A. (1789). *Diccionario da Lingua Portugueza Composto pelo Padre D. Rafael Bluteau, reformado, e acrescentado por Antonio de Moraes Silva*. Ferreira.

Print Dictionary Entries

Corominas, J. & Pascoal, J.A. (1980-1983). Armada. In *Diccionario crítico etimológico castellano e hispánico*. Gredos. (Vol. 1, p. 22).

Corominas, J. & Pascoal, J.A. (1980-1983). Escuela. In *Diccionario crítico etimológico castellano e hispánico*. Gredos. (Vol. 2, pp. 714-715).

Corominas, J. & Pascoal, J.A. (1980-1983). Falso. In *Diccionario crítico etimológico castellano e hispánico*. Gredos. (Vol. 2, p. 841).

Corominas, J. & Pascoal, J.A. (1980-1983). Mesa. In *Diccionario crítico etimológico castellano e hispánico*. Gredos. (Vol. 4, p. 57).

Corominas, J. & Pascoal, J.A. (1980-1983). Muñeca. In *Diccionario crítico etimológico castellano e hispánico*. Gredos. (Vol. 4, pp. 188-190).

- Corominas, J. & Pascoal, J.A. (1980-1983). Saco. In *Diccionario crítico etimológico castellano e hispánico*. Gredos. (Vol. 5, p. 121-122).
- Houaiss, A. (2003). Armada. In *Dicionário Houaiss da língua portuguesa*. Temas e Debates. (Vol. 1, pp. 375-376).
- Houaiss, A. (2003). Boneca. In *Dicionário Houaiss da língua portuguesa*. Temas e Debates. (Vol. 1, p. 628).
- Houaiss, A. (2003). Escola. In *Dicionário Houaiss da língua portuguesa*. Temas e Debates. (Vol. 2, p. 1563).
- Houaiss, A. (2003). Falso. In *Dicionário Houaiss da língua portuguesa*. Temas e Debates. (Vol. 2, p. 1691).
- Houaiss, A. (2003). Mesa. In *Dicionário Houaiss da língua portuguesa*. Temas e Debates. (Vol. 3, pp. 2465-2466).
- Houaiss, A. (2003). Saco. In *Dicionário Houaiss da língua portuguesa*. Temas e Debates. (Vol. 3, p. 3226).
- Machado, J. P. (1967). Armada. In *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência. (Vol. 1, p. 307).
- Machado, J. P. (1967). Boneca. In *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência. (Vol. 1, p. 448).
- Machado, J. P. (1967). Escola. In *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência. (Vol. 2, p. 923).
- Machado, J. P. (1967). Falso. In *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência. (Vol. 2, p. 1010).
- Machado, J. P. (1967). Mesa. In *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência. (Vol. 2, p. 1554).
- Machado, J. P. (1967). Saco. In *Dicionário etimológico da língua portuguesa : com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. Confluência. (Vol. 3, p. 2046).

Morais, A. (1789). Armada. In *Diccionario da Lingua Portuguesa Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Ferreira (Vol. 1, p. 112).

Morais, A. (1789). Boneca. In *Diccionario da Lingua Portuguesa Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Ferreira (Vol. 1, p. 189).

Morais, A. (1789). Escola. In *Diccionario da Lingua Portuguesa Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Ferreira (Vol. 1, p. 533).

Morais, A. (1789). Falso. In *Diccionario da Lingua Portuguesa Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Ferreira (Vol. 1, p. 596).

Morais, A. (1789). Mesa. In *Diccionario da Lingua Portuguesa Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Ferreira (Vol. 2, p. 77).

Morais, A. (1789). Saco. In *Diccionario da Lingua Portuguesa Composto pelo Padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva*. Ferreira (Vol. 2, p. 364).

Online Dictionary Entries

Badan Pengembangan dan Pembinaan Bahasa. (2016) Armada. In *KBBI Daring*. Retrieved January 24, 2022, from <https://kbbi.kemdikbud.go.id/entri/armada>

Badan Pengembangan dan Pembinaan Bahasa. (2016) Tempo. In *KBBI Daring*. Retrieved January 31, 2022, from <https://kbbi.kemdikbud.go.id/entri/jatuh%20tempo>

Porto Editora. (n.d.) Armada. In *Infopédia*. Retrieved January 4, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/armada>

Porto Editora. Aula. In *Infopédia*. Retrieved January 25, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/aula>

Porto Editora. (n.d.) Legenda. In *Infopédia*. Retrieved February 2, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/legenda>

Porto Editora. (n.d.) Lenda. In *Infopédia*. Retrieved February 2, 2022, from <https://www.infopedia.pt/dicionarios/lingua-portuguesa/lenda>

Online Resources

Ardanawati, I. (2019, June 25). 25 Juni 1938, Sejarah Kongres Bahasa Indonesia I: Meresmikan Bahasa Persatuan. *Tirto.id*. Retrieved on March 26, 2022 from <https://tirto.id/sejarah-kongres-bahasa-indonesia-i-meresmikan-bahasa-persatuan-ecjH>

Bidang Pengembangan, Pusat Pengembangan dan Perlindungan. (n.d.). *Bagaimana Sebuah Kata Masuk ke KBBI*. Retrieved on January 21, 2022 from <https://badanbahasa.kemdikbud.go.id/lamanbahasa/artikel/2547/bagaimana-sebuah-kata-masuk-ke-kbbi/>

Haspelmath, M. & Tadmor, U. (eds.) (2009) World Loanword Database. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved October 10, 2021 from <http://wold.clld.org>

Indonesia.Go.Id. Portal Informasi Indonesia. (n.d.). *Agama*. Retrieved on January 24, 2022 from <https://www.indonesia.go.id/profil/agama>

United Language Group. (n.d.). *Linguistic Code-Switching: What It Is and Why It Happens*. Retrieved on March 26, 2022 from <https://www.unitedlanguagegroup.com/blog/linguistic-code-switching>

Wikipedia Entries

Indonesian language. (2022, February 14). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Indonesian_language&oldid=1070887199

Indo-Portuguese creoles (2022, January 19). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Indo-Portuguese_creoles&oldid=1063533786

Islam in Indonesia. (2022, January 24). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Islam_in_Indonesia&oldid=1066621470

Javanese language. (2022, January 26). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Javanese_language&oldid=1067983437

Kawi language. (2022, January 22). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Kawi_language&oldid=1061636037

List of islands by population. (2022, January 13). In *Wikipedia*.
https://en.wikipedia.org/w/index.php?title=List_of_islands_by_population&oldid=1065505979

Mardijker Creole (2022, January 19). In *Wikipedia*.
https://en.wikipedia.org/w/index.php?title=Mardijker_Creole&oldid=1053076517

Stratum (linguistics). (2022, February 3). In *Wikipedia*.
[https://en.wikipedia.org/w/index.php?title=Stratum_\(linguistics\)&oldid=1069609575](https://en.wikipedia.org/w/index.php?title=Stratum_(linguistics)&oldid=1069609575)

Youth Pledge. (2021, November 16). In *Wikipedia*.
https://en.wikipedia.org/w/index.php?title=Youth_Pledge&oldid=1055590381

Appendix I

List of known Portuguese Loanwords in Indonesian Language

No.	Indonesian	Portuguese Origin	English equivalent	Frequency Rank
1.	adven	advento	advent	129,995
2.	algojo	algoz	executioner	30,009
3.	almari (lemari)	armário	wardrobe	34,459
4.	altar	altar	altar	21,750
5.	antero	inteiro	whole	122,973
6.	arena	arena	arena	4,881
7.	arloji	relógio	watch/clock	37,816
8.	armada	armada	armada	5,150
9.	aula	aula	big, multifunctional room	8,841
10.	baluarti	baluarte	stronghold	1,135,748
11.	bangku	banco	bank (to sit on)	4,161
12.	banyo	banjo	banjo	752,441
13.	bantal	avental	pillow	10,682
14.	Belanda	Holanda	Holland	967
15.	baret	bareta	beret	37,000
16.	beludru	veludo	velvet	54,148
17.	bendera	bandeira	flag	3,671
18.	beranda	veranda	veranda	24,906
19.	biola	viola	violin	22,090
20.	bola	bola	ball	918
21.	bolu	bolo	cake	44,015
22.	boneka	boneca	doll	6,584
22.	botol	botelha	bottle	3,463
23.	cerutu	charuto	cigar	40,770
24.	cokelat	chocolate	chocolate	5,558
25.	dadu	dado	dice	30,388
26.	dansa	dançar	dance	28,119
27.	duda	viúva	widower	32,227
28.	gagu	gago	stutterer	139,080
29.	gardu	guarda	guard	17,555
30.	garpu	garfo	fork	29,071
31.	gereja	igreja	church	2,010
32.	gitar	guitarra	guitar	6,080
33.	Inggris	inglês	English	458
34.	jendela	janela	window	3,570

35.	jurnal	jornal	journal	4,085
36.	kaldera	caldeira	volcano crater	53,988
37.	kaldu	caldo	broth	24,986
38.	kamar	câmara	room/chamber	406
39.	kampung	campo	countryside/village	1,908
40.	kanon	canon	cannon	41,851
41.	kantin	cantina	canteen	11,556
42.	kapten	capitão	captain	5,553
43.	karambol	carambola	fiddle	133,956
44.	kartu	cartão	card	1,105
45.	kastel	castelo	castle	34,671
46.	keju	queijo	cheese	10,044
47.	kemeja	camisa	shirt	10,848
48.	komedi	comédia	comedy	12,140
49.	kompeni	companhia	company	55,042
50.	kereta	carreta	cart/train	1,570
51.	koral	coral	coral	35,382
52.	kubis	couves	cabbage	28,291
53.	kursus	cursos	courses	3,864
54.	kutang	cotão	undershirt	88,104
55.	labu	nabo	pumpkin	19,354
56.	lampion	lâmpião	lantern	27,909
57.	legenda	legenda	legend	7,518
58.	lelang	leilão	bidding	3,513
59.	lemon/limau	limão	lemon	16,041
60.	lentera	lanterna	lantern	42,043
61.	lenso	lenço	tissue	270,556
62.	loji	loja	shop	92,979
63.	mandor	mandador	foreman	30,480
64.	marakas	maracas	maracas	313,425
65.	markisa	maracujá	passionfruit	49,689
66.	marmot	marmota	guinea pig	145,477
67.	martil	martelo	hammer	60,870
68.	martir	mártir	martyr	32,599
69.	meja	mesa	table/desk	1,982
70.	mentega	manteiga	butter	17,287
71.	meski (pun)	mas que	though	956
72.	Minggu	domingo	Sunday	790
73.	misa	missa	misa	17,397
74.	nanas	ananás	pineapple	16,371
75.	Natal	Natal	Christmas	3,054

76.	nyonya	senhora	missus, lady	3,054
77.	noda	nódoa	stain	11,565
78.	nona	dona	miss	26,212
79.	padri	padre	priest	198,327
80.	palsu	falso	false, fake	2,995
81.	Paskah	Páscoa	Easter	10,100
82.	pastel	pastel	pie	33,061
83.	pegang	pegar	to hold	11,962
84.	peluru	pelouro	bullet	6,319
85.	pena	pena	pen	16,142
86.	peniti	alfinete	safety pin	57,556
87.	pepaya	papaia	papaya	12,057
88.	persero	parceiro	partner, stakeholder	4,244
89.	pesiar	passrear	Cruise	15,452
90.	pesta	festa	party	2,613
91.	Prancis	francês	french	3,014
92.	prangko	franco	stamp	35,954
93.	pigura	figura	frame	54,226
94.	pion	peão (xadrez)	pawn (chess)	75,110
95.	pita	fita	ribbon	7,415
96.	Portugis	português	Portuguese	11,873
97.	puisi	poesia	poem	3,391
98.	rantai	corrente	chain	5,080
99.	renda	renda	lace	45,389
100.	roda	roda	wheel	2,328
101.	ronda	ronda	night patrol	34,024
102.	rosario	rosario	rosary	53,406
103.	Sabtu	sábado	Saturday	888
104.	sabun	sabão	soap	6,621
105.	saku	saco	pocket	8,669
106.	salto	salto	jump	50,526
107.	seka	secar	to dry	85,957
108.	sekolah	escola	school	215
109.	selada	salada	salad	34,049
110.	selai	geleia	jam	24,568
111.	sepatu	sapato	shoe	2,549
112.	serdadu	soldado	soldier	18,423
113.	silet	gilete	razor	52,765
114.	sinyo	sinhô/senhor	young man of European origin	252,412
115.	sita	citar (confiscar)	To confiscate	22,531

116.	tanjidor	tangedor	a traditional musical style	99,766
117.	tapioka	tapioca	tapioca	32,382
118.	teledor	traidor (negligente)	careless	60,355
119.	tembakau	tabaco	tabacco	5,550
120.	tempo	tempo	time/speed	4,414
121.	tenda	tenda	tent	6,682
122.	terigu	trigo	wheat flour	10,521
123.	tinta	tinta	ink	9,549
124.	tolol	tolo	stupid	33,860
125.	tuala	toalha	towel	84,433
126.	tukar	trocar	to exchange	3,061