

## Balinese cultural representation from the perspective of the Balinese wordlist

I Made Sena Darmasetiyawan

Universitas Udayana, Indonesia

\*Corresponding Author. Email: [sena\\_darmasetiyawan@unud.ac.id](mailto:sena_darmasetiyawan@unud.ac.id)

### Abstract

Based on the data gathered from the PARADISEC catalogue of the Balinese wordlist, as part of the OCSEAN project, 1,228 words were recorded and classified into several categories, such as the body, agriculture and vegetation, animals, the physical world, basic actions and technology, food and drink, spatial relations, kinship, motion, sense perception, quantity, emotions and values, cognition, time, speech and language, the house, warfare and hunting, possession, clothing and grooming, religion and belief, social and political relations, law, as well as miscellaneous function words. This study examines the Balinese wordlist from the OCSEAN data and explores the correlation between the number of words gathered and the speakers' language recognition and comprehension within their community. Since the process of word recognition involves identifying both meaning and form (orthographic and phonological), these aspects originating from each speaker are generally influenced by their surrounding community. Through mapping the quantitative data, possible borrowings, and accuracy in the wordlist, several assumptions can be made. For example, the high number and accuracy of words related to agriculture (including its associated resources) may be attributed to the fact that many Balinese speakers come from farming communities. The implications of this study suggest that future research may benefit from combining pre-existing linguistic data with the psychological aspects of speakers and the cultural characteristics of their communities.

### Article History

Received:

5 June 2025

Accepted:

16 November 2025

Available online:

30 November 2025

### Keywords

Balinese wordlist;  
word recognition;  
community  
influence; lexical  
categorization;  
linguistic  
anthropology

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



**Citation (APA Style):** Darmasetiyawan, I. M. S. (2025). Balinese cultural representation from the perspective of the Balinese wordlist. *LingTera*, 12(2), 118–126. <https://doi.org/10.21831/lt.v12i2.86302>

## INTRODUCTION

In the conceptual order, prior to the process of language acquisition, there were language perception, language representation, and language recognition processes (Darmasetiyawan et al, 2024). Distinction to these concepts is quite thin (it often overlaps with each other), which points out to the most apparent process in language recognition that involves the access to mental lexicon. Recognition process involved segmental and suprasegmental content that is used to identify word boundaries and distinction between phonological and conceptual representation (Mcqueen in Gaskell, 2011). In the orthographical level, this can be identified with segmentation (i.e., syllables, phrases, clauses), while phonological level can be identified through articulatory rules and its semantic lexicon. Here, recognition process would include the ability to interpret and give meaning from the speaker. In order to see any discrepancy of words to be listed in the effort of constructing wordlist, the state and role of each native speaker that act as the data source, will have to be considered psychologically. The fundamental reason behind choosing any of the native speaker will have to rely on that person's word recognition level as well (this level generally went in similar fashion with the language frequency of exposure; hence, the speaker age). It does important to bear



in mind though, there are ceilings to this age variable, since one cannot guarantee that given the old age of the speaker, chances of few conditions to language impairment will start to show. The use of video and audio recording during wordlist data collection from fieldwork research can be considered as a reliable technique, since a sufficient control can be applied. Erratic choice of words by the recorded speaker as a result of recording pressure, or the identification of small errors in the way linguist captured the written words when there was no pressure, can sufficiently identify the participant state during the process of data collection; thus, producing a reliable data (Klamer et al, 2021). This technique, however, required the gathering of sufficient number of locals (native speakers) in order to reach a consensus for the translational equivalent of each word that will be listed.

Some variables screening that includes stable social context and fluent native speaker might seem to suffice the categorization, but the sufficient number to reach this consensus can still create a significant gap into the methodological foundation, since some questions related to these variables persist; i.e., how fluent should it be or how to count the consensus statistically (since it represents the whole community). In the case of this wordlist, it already includes the significance of crosslinguistic (multilingual) standing of each native speaker, since similarities between related language helps the identification and control for contingent features that are not obvious when only one language is investigated (Pye, 2017). Nevertheless, there were bound to be several shortcomings that may have been unnoticed; i.e., how during production studies, failure might come from preference and memory or processing limitation (Ambridge and Lieven, 2011).

In the effort to pursue the correlation of all words in Balinese wordlist with the characteristics of Balinese community, a different perspective to word recognition process would be required. Sufficient information on the historical background of the community would have to be gathered, in order to legitimize the foundation to the speaker word recognition that can be reflected from the accuracy of each word in the wordlist. The present study will explain how Balinese word form and meaning can reflect the characteristics of Balinese community through the lens of linguistics (particularly wordlist). The following result and discussion section in this article will first outline results from all words translation equivalent accuracy, and then combine it with sociological and historical perspectives of Balinese community. It will, therefore, lay out a culturally interpretative analysis that can suggest future researches to move beyond existing wordlist such as WOLD or Swadesh.

## RESEARCH METHOD

This study was initially brought up by the general concept of linguistics is how similar features in languages can be seen from the probability of common ancestors from the speakers, mutual contact, and adopted features amongst them (Klamer and Moro, 2023). Originally, the notion of 100-item list was introduced to describe the notion of basic vocabulary that can be used for comparative linguistic studies (Swadesh in Tadmor et al, 2012). One major project that significantly impacted the use of wordlist is Loanword Typology (LWT) from Haspelmath and Tadmor (2009) that used fixed list of lexical meaning of 1460 words. On the common foundation that body part terms are unlikely to be borrowed – new artifacts are often borrowed, this study also counted the possible borrowing words from similar meaning list. It is important to acknowledge that this list of words is not automatically a wordlist, per say, since the list is originally a list of meaning that could be (probability) relevant in any language, not words from that language. As mentioned by Haspelmath and Tadmor (2009), LWT meaning list includes 1310 meaning list from Intercontinental Dictionary Series (IDS), with 150 additional meanings from three categories of concept important to geographical regions beyond the IDS, meanings that appear on the 207 Swadesh list but not in the IDS, and common meanings pertaining to modern life.

In Edwards (2021) study, it was clearly stated that while we can assume the semantics of certain words are similar with words in another language, we cannot assume that they are identical – it may well have additional meaning that are not present in the other language form. In the cases of 3 or more similar words, as Edwards (2021) referred to out-comparison, there is a possibility that these correspondences indicated loans/borrowings, rather than the result of shared inheritance. The

argument of using possible borrowings strictly in one semantic type in this study, so that it can be clearly distinguished with homonyms that might appear in Indonesian; i.e., pasang of high tide, pasang of pair, and pasang of wear (Haspelmath and Tadmor, 2009). In Tadmor et al (2012) study, it was clearly mentioned that it has long been known, how languages are more likely to borrow nouns than verbs, as well as adjectives and adverbs. In relation to the concept of language contact, possible borrowings, or in this case, lexical borrowing, can be seen in Table 1.

**Table 1. Lexical Borrowing in Thomason’s Borrowing Scale  
(Thomason in Klammer and Moro, 2023)**

No.	Intensity of contact	Type of speakers	Borrowed elements
1.	Casual	Few bilinguals among borrowing-language speakers, borrowers need not be fluent in the source language.	Only non-basic vocabulary. Only content words; most often nouns, verbs, adjectives, and adverbs.
2.	Slightly more intense	More fluent bilinguals among borrowing-language speakers, but they are probably still a minority.	Still non-basic vocabulary. Function words (e.g., conjunctions and adverbial particles like ‘then’) as well as content words.
3.	More intense	A conspicuous number of bilinguals among borrowing-language speakers, attitudes and other social factors favoring borrowing.	Basic and non-basic vocabulary. More function words, including closed-class items as pronouns and low numerals; derivational affixes.
4.	Intense	Very extensive bilingualism among borrowing-language speakers, social factors strongly favoring borrowing.	Heavy lexical borrowing in all sections of the lexicon.

Balinese speakers are bilinguals to the extent that they also have to master Indonesian as their national language (Darmasetiyawan and Ambridge, 2023). While Balinese speakers in general can be categorized as having slightly more intense language contact through the lens of bilingual position of Balinese and Indonesian language that has been introduced to the speakers from their early age, it is important to also emphasize on the gradation of this effect. Different exposure to other language and education level can generate different level of language contact; speakers in high density population (i.e., major city such as Denpasar) may be considered having intense language contact, while the opposite applied to Balinese speakers in remote villages or mountain areas (i.e., around Karangasem region). Therefore, the need to explore other possible semantic differences would be necessary to grasp specific effects portrayed from the Balinese culture to its wordlist. In terms of borrowing, Balinese language was compared to Indonesian language (and was marked as borrowing in terms of orthographical or phonological similarity), as it is the language that can greatly impacted language production of Balinese speakers.

In terms of using wordlist, or commonly considered as lexicostatistic approach, several studies conducted in Indonesia (i.e., Sulistyono and Fernandez, 2015; Ntelu and Djou, 2017; Lailiyah and Wijayanti, 2022; Rahman and Agus, 2022), was primarily done using Swadesh list to figure out relationship (kinship) amongst regional languages. Using the same approach, Kaplan (2017) also debated the nature of Swadesh list that despite the trend to move onto electronic computing, historical linguistics perspective to the comparative linguistics actually suggest the use of basic wordlist to enforce future changes within the wordlist. This study was originated from the claim of Heggarty (2010), where he stated that language relationship from wordlist comparison would benefit more by exploring emerging patterns from meanings.

Based on Basa Bali Wiki; <https://dictionary.basabali.org/> there are 6 different classifications of Balinese speech levels, generally divided into two, high and low register. Three low register speech levels are *Mider*, *Kasar*, and *Andap*, while the other three for high register are *Alus Sor*, *Alus Mider*, and *Alus Singgih*. There were, however, talk of how the Balinese speech level were classified into four, only one low register (or *Kepara*), and the same three for its high register. In Balinese language, in terms of the Principle of Pragmatic Consistency used by Arka (2005), he mentioned that there is a grammatical inconsistency of the traditional speech level of Balinese (high and low, or *alus* and *lumrah*), from predicate alternatives to the potential number of word forms, social-relation

restrictions in the arguments, and constraint on the addressee. It was clear that due to this significant difference on level of fluency and comprehension of Balinese speakers, data taken in the wordlist were based upon the most common level of speech – which was low register (*Andap/Kepara*). The use of 24 semantic fields in Balinese language corresponded to the same semantic fields used in World Loanword Database (WOLD), which can be accessed through <https://wold.clld.org/meanings>. The Balinese wordlist of 1228 words that classified into 23 semantic fields can be viewed from in OSF (<https://osf.io/89mzg/overview>). It is important to note that although the data was taken from informant during data collection to fill the wordlist, certain aspects of the data collection process cannot be disclosed in this article, as doing so could enable the dataset to be used for purposes beyond the scope of the present study.

## RESULTS AND DISCUSSION

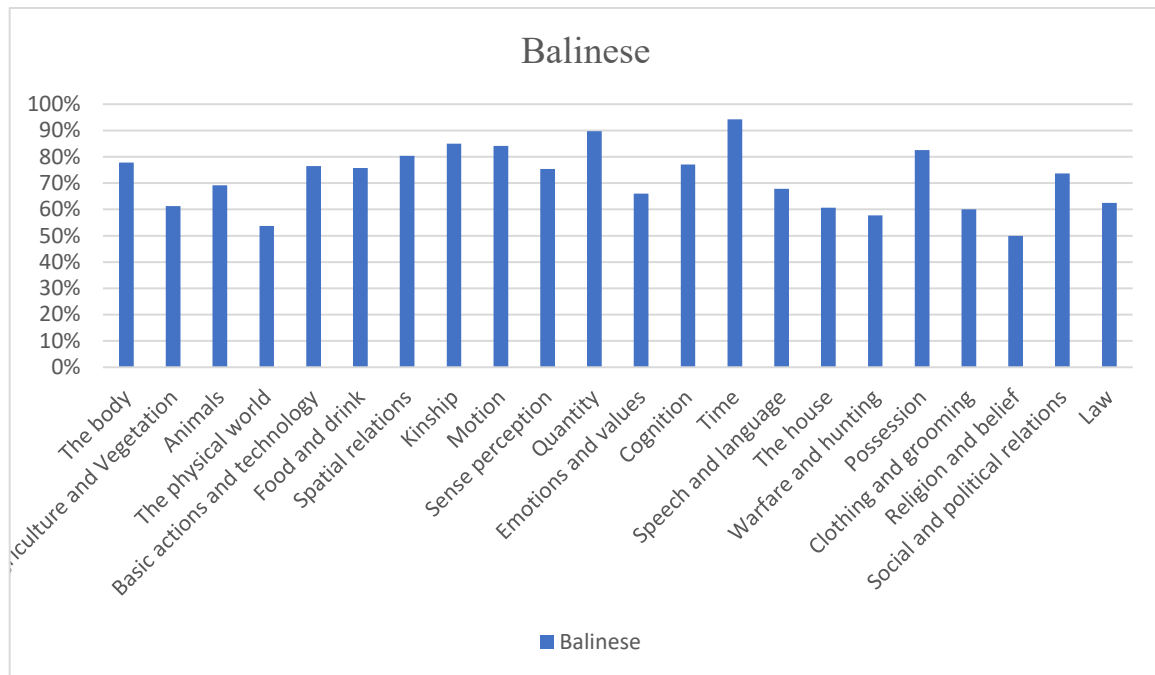
Based on a general view of the wordlist, the table that can be seen below is the number of wordlist when deducted by possible borrowings and the case of replication; the use of multiple words (semantically and orthographically the same; exactly the same word that does not differ in meaning) in that same semantic field of Balinese; i.e., the Balinese word *siyap* in animal is used twice for fowl and chicken. This can be observed in Table 2.

**Table 2. Balinese Wordlist Deducted by Possible Borrowings and Replication**

No	Semantic Fields	Number	Borrowing	Replicate	Amount
1	The body	189	34	8	147
2	Agriculture and Vegetation	111	31	12	68
3	Animals	104	29	3	72
4	The physical world	93	36	7	50
5	Basic actions and technology	81	17	2	62
6	Food and drink	66	14	2	50
7	Spatial relations	61	8	4	49
8	Kinship	60	3	6	51
9	Motion	57	7	2	48
10	Sense perception	57	13	1	43
11	Quantity	49	4	1	44
12	Emotions and values	47	13	3	31
13	Cognition	35	7	1	27
14	Time	35	2	0	33
15	Speech and language	28	7	2	19
16	The house	28	9	2	17
17	Warfare and hunting	26	8	3	15
18	Possession	23	3	1	19
19	Clothing and grooming	20	8	0	12
20	Religion and belief	20	10	0	10
21	Social and political relations	19	4	1	14
22	Law	8	3	0	5

Based on the list above and a comparison to the LWT, there were two semantic field that was not used; it is miscellaneous function words and modern world. Modern world that contains words such as television, car, or radio, is arguably neglected due to the relevance of them with how Balinese (or other similar languages in Austronesia) that does not have similar concept back when this language was originated – possible representation of these words would only appear in borrowings – therefore, resulted in no actual data for the Balinese wordlist itself. In fact, it was not

until 1962 that Indonesian State Television (TVRI) begun broadcasting in Bali (Kitley, 2000). The opposite case happened to miscellaneous function words that can be omitted due to no borrowings occurred. When deducted by possible borrowings and duplications, the percentage of Balinese wordlist can be seen in Figure 1.



**Figure 1. Balinese Wordlist Number Deducted by Possible Borrowings and Replicate**

The argument to subtract the number of words in each semantic field and viewed in percentage, refers to the notion that by omitting borrowings and replication, the number of each word (or its semantics) can be counted to clearly represent its specific (unique) meaning. The high percentage of semantics in time and quantity represents how Balinese speakers has a clear concept of time and quantity, as both can be related to the account of events and the significance of transactions. Historically, there were past events that still recorded in detail through royalty families in Bali and how arguably Bali put significant emphasize in its trading and commerce (as reflected in quantity). Another clear example to emphasize quantity amongst Balinese is how they have their own calendar system that can accurately schedule proper time to hold religious ceremony (i.e., *dewasa rainan*), which was based upon Pawukon cycle (Chatterjee, 1997).

The opposite can also be seen from religion and belief, as well as the physical world semantic fields. Their low percentage of semantic wordlist was arguably due to the movement of population (that brought along naming concepts of surroundings) and religion (mostly Hindu) that came from the island of Java. In terms of religion, other religions moved along with the population (sometimes assimilated; i.e., Hindu and Christianity in Palasari village, or Hindu and Moslem in Loloan village). This also brought changes to Hindu religion itself that is significantly assimilated to the Balinese culture, other Hindu (as religion) in other places aside from Bali might have significant differences in terms of practices or rituals. In terms of physical world, the low percentage can be taken as a fact due to how most words in modern world is also borrowing – this clearly showed the impact of modern world also brought to physical world.

It is also worth noting that when compared to the previous table, despite having the largest number of words to represent its semantic field, the body did not necessarily bring a lot of borrowings as well, the exact number is still fewer than in the physical world. Further comparison of this finding to the global perspective can be seen by looking at data from Tadmor et al (2012) and WOLD.

Semantic Fields	Percentage		Semantic Fields	Percentage
Religion and belief	50%		Religion and belief	41%
Clothing and grooming	40%		Clothing and grooming	39%
The physical world	39%		The house	37%
Law	38%		Law	34%
The house	32%		Social and political relations	31%
Warfare and hunting	31%		Agriculture and Vegetation	30%
Agriculture and Vegetation	28%		Food and drink	29%
Animals	28%		Warfare and hunting	28%
Emotions and values	28%		Possession	27%
Speech and language	25%		Animals	26%
Sense perception	23%		Cognition	24%
Food and drink	21%		Basic actions and technology	24%
Social and political relations	21%		Time	23%
Basic actions and technology	21%		Speech and language	22%
Cognition	20%		Quantity	21%
The body	18%		Emotions and values	20%
Spatial relations	13%		The physical world	20%
Possession	13%		Motion	17%
Motion	12%		Kinship	15%
Quantity	8%		The body	14%
Time	6%		Spatial relations	14%
Kinship	5%		Sense perception	11%

**Figure 2. Comparison of Balinese Borrowings with Data from Tadmor et al. (2012) and WOLD (<https://Wold.Cild.Org>)**

The data of semantic fields taken from Tadmor et al (2012) and WOLD, showed that compared to Balinese wordlist, three of the most borrowings occurred in the same semantic fields of religion and belief, clothing and grooming, as well as law. The same rank from these three semantic fields showed that what occurred to Balinese people in terms of religion and belief, clothing and grooming, as well as law also occurred globally – marked how global movement and changes also occurred in Bali. In terms of significant changes, however, low percentage in the global physical world that compared with high percentage in the Balinese borrowing, further solidify the argument that Balinese has significantly affected by foreign (global) influence of modern world semantics. In other hand, low percentage of social and political relations, as well as possession, when compared to data from Tadmor et al (2012) and WOLD, showed that despite the strong foreign influence, emphasize to the community values in social and political realm, along with huge respect to possession and cultural practices in specific family group (as well as caste systems) reflected how the semantics in both fields were still mostly maintained. This caste system would also include marriage, residence, and social status (Geertz and Geertz, 1975).

From Edwards (2021) study, it is important to notice and compare the case of multiple reflexes in the wordlist, where morphologically related forms (i.e., derivatives) in a single construction showed indication of semantic differences, which in turn created the notion of how duplicates was used by the speakers to reflect semantic or orthographic relatedness between words; marked with partial use in a phrase (usually a single word) to express phrases; i.e., the word *nasi* in food and drink semantic field was used in *nasi goreng* (fried rice) and *goh nasi* (viand) as it both contain huge proportion of information *nasi* (cooked rice).

The use of this phrase construction creates anaphoric effect that can be viewed to see other relevant analysis on the origin of these words. Another example that can be seen in the wordlist, is the word bone – *tulang* (tu.laŋ) that can be categorized as duplication, due to other entries of clavicle – *tulang baong* (tu.laŋ.ba.ɔŋ), spine – *tulang tundun* (tu.laŋ.tun.dɔn), and rib – *tulang iga* (tu.laŋ.i.ga). These instances were deducted from the wordlist count (and displayed in percentage), since the other entries did not represent a single meaning on its own – rather, an expansion of the same word that



tried to specify the meaning with additional words. Moreover, there are instances, where some of the duplications were also a case of borrowing; i.e., while the word *tulang* is arguably not a borrowing, the word *iga* that accompanies it, can be categorized as borrowing. Hence, the initial wordlist is not deducted by instances of duplications, since the expansion of meaning did not necessarily reflect same meaning (as in replication). By definition, replication is a term rarely used, due to the nature of borrowing from one language to another, which creates a vague distinction to this term – not necessarily loan, which can also be defined as re-use of re-creation of grammatical pattern (Wiemer, Wälchli, and Hansen, 2012). In terms of lexical bases, the closest accepted replication to this phenomenon is the borrowing case of morphological formatives by Gardani (2021) in the use of affixes and abstract morphological patterns. However, these three terms (borrowing, replicate, and duplicate) were all used in the study to put a clear boundary in lexical instances that can reflect cultural interpretation from national-regional languages relationship for the speakers. In short, borrowing is occurrence where the word clearly has other equivalence in the target language, replicate refer to the use of same word to express different meanings, and duplicate refer to the use of partially same word to express different meanings.

**Table 3. Comparison with the Number of Duplicates in Balinese Wordlist**

No.	Semantic Fields	Borrowing	Replicate	Duplicate
1	Quantity	4	1	12
2	The body	34	8	11
3	The physical world	36	7	8
4	Kinship	3	6	8
5	Animals	29	3	7
6	Spatial relations	8	4	7
7	Agriculture and Vegetation	31	12	6
8	Cognition	7	1	4
9	Sense perception	13	1	3
10	Basic actions and technology	17	2	2
11	Time	2	0	2
12	Food and drink	14	2	1

Table 3 showed that from 22 semantic fields taken account in the Balinese borrowings, only 12 that have instances of duplication. While the number of replications can indicate how narrow semantics in that field can be (since one word can represent more than one entry), the number of duplications, however, indicate that there were expansions to the semantic of that word to be embedded in other entries. Through the table, duplications were sorted from the highest one in quantity semantic field to the lowest one in food and drink semantic field. Therefore, it can be seen here that despite having one of the lowest percentages of borrowing, the number of duplications occurred in quantity semantic field was quite large, indicating that most of the entries requires expansions of meaning from its few core entries; i.e., ten – *dasa* (da.sə) and twenty – *duang dasa* (du.aŋ.da.sa) or ninety – *sie dasa* (si.ja.da.sa). The same pattern also occurred in other language (i.e., Indonesian), where there were no unique entries for numbers in multiplication (i.e., tens or hundreds), such as two – *dua* just added with the word *puluh* when in tens form; twenty – *dua puluh*. This instance did not occur in time, as the significance of this semantic field further emphasized by how it has low percentage of borrowings (with no replication at all) and also low number of duplications. Although several assumptions can be made, but this table can generally reflect which semantic fields emphasized in the Balinese community (as it persisted in the Balinese speakers' comprehension), as well as how it was able to be maintained longer than the others. This case, particularly this example in numbering, is not unique to Balinese, as in Javanese, for example, the vocabulary for 21 would be *selikur*, 22 would be *rolikur*, 23 would be *telulikur*, and so on; the use of *likur* are repetitive to mark the twenty. Therefore, several instances cannot be taken out as representation of a distinctive

Balinese morphological pattern. Quantity wise, however, can reflect cultural representation (albeit similar to other regional languages) when related to the Indonesian language as national language. It was worth pointing out that since this approach is relatively new in cultural interpretation, there were no study in other regional languages yet.

In terms of percentage comparison, the Balinese wordlist that was deducted by borrowings and replications, compared to its percentage of duplication, similar relationship can be seen – two largest (in percentage) duplications occurred in quantity and kinship semantic fields, despite amongst one of the lowest in terms of borrowings and replications. The pattern that can be seen here is how the percentage of replication and duplication completes each other – low number appear in replication did not necessarily mean low number in duplication as well. This pattern suggested that when certain expansion to meaning occurred in a specific semantic field, it would also imply that most of the words in that wordlist were well represented (moderately low or very low replication). Furthermore, further justification to previous claim of how Balinese emphasize on words in quantity and kinship semantic field (to justify the low number of borrowings and replication), can also be supported by the expansion of meaning (hence, duplication), which imply how the speakers needed additional specific terms due to the intensity or frequency and importance of the semantic field itself in their daily life.

## CONCLUSION

Although the data was taken from a portion of the OCSEAN project, a preliminary and rather manual approach can provide quick overlay on the accuracy of the data. Generally, Balinese has significant percentage of combined borrowing, replication, and duplication in the semantic fields that involves time and quantity. The comparison to data from Tadmor et al (2012) and WOLD showed that while Balinese has similar percentage in terms of religion and belief, clothing and grooming, as well as law, striking differences can be found in terms of the physical world, social and political relation, as well as possession. Further study should try to elaborate and use computing models that can help illustrate the length of duplication in stretching or narrowing meaning of specific words.

## ACKNOWLEDGEMENT

This study was a part of OCSEAN project that received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.873207.

## REFERENCES

- Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511975073>
- Arka, I. W. (2005). Speech levels, social predicates and pragmatic structure in Balinese: A lexical approach. *Pragmatics*, 15(2–3), 169–203.
- Atai, M. R., & Habibi, P. (2009). Exploring sub-disciplinary variations and generic structure of applied linguistics research article introductions using CARS model. *Applied Linguistics*, 2, 26–51.
- Chatterjee, S. K. (1997). Balinese traditional calendar. *Indian Journal of History of Science*, 32(4), 325–347.
- Darmasetiyawan, I. M. S., & Ambridge, B. (2022). Syntactic representations contain semantic information: Evidence from Balinese passives. *Collabra: Psychology*, 8(1), Article 33133. <https://doi.org/10.1525/collabra.33133>
- Darmasetiyawan, I. M. S., Aryawibawa, I. N., & Mulyawan, I. W. (2024). *Psycholinguistics*. Mega Press Nusantara.
- Edwards, O. (2021). *Rote-Meto comparative dictionary*. ANU Press.
- Gardani, F. (2021). On how morphology spreads. *Word Structure*, 14(2), 129–147. <https://doi.org/10.3366/word.2021.0184>
- Gaskell, M. G. (2011). *The Oxford handbook of psycholinguistics*. Oxford University Press.
- Geertz, C., & Geertz, H. (1975). *Kinship in Bali*. University of Chicago Press.



- Haspelmath, M., & Tadmor, U. (2009). *Loanwords in the world's languages: A comparative handbook*. De Gruyter Mouton. <https://doi.org/10.1515/9783110218442>
- Heggarty, P. (2010). Beyond lexicostatistics: How to get more out of 'word list' comparisons. *Diachronica*, 27(2), 301–324. <https://doi.org/10.1075/dia.27.2.07heg>
- Kaplan, J. (2017). From lexicostatistics to lexomics: Basic vocabulary and the study of language prehistory. *Osiris*, 32(1), 202–223. <https://doi.org/10.1086/694093>
- Kitley, P. (2000). *Television, nation, and culture in Indonesia*. Ohio University Press.
- Klamer, M., Edwards, O., Fricke, H., Gialitaki, Z., Moro, F., Palmér, A., Saad, G., Sulistyono, Y., Visser, E., & Wu, J. (2021). Practicalities of language data collection and management in and around Indonesia. *Wacana: Journal of the Humanities of Indonesia*, 22(2), 467–521.
- Klamer, M., & Moro, F. (Eds.). (2023). *Traces of contact in the lexicon*. Brill. <https://doi.org/10.1163/9789004529458>
- Lailiyah, N., & Wijayanti, F. I. (2022). Kekerabatan bahasa Jawa, Bali dan Bima: Perspektif linguistik historis komparatif. *Linguistik Indonesia*, 40(2), 327–345. <https://doi.org/10.26499/li.v40i2.308>
- Ntelu, A., & Djou, D. N. (2017). The language family relation of local languages in Gorontalo Province: A lexicostatistic study. *Journal of Arts & Humanities*, 6(11), 48–59. <https://doi.org/10.18533/journal.v6i11.1285>
- Pye, C. (2017). *The comparative method of language acquisition research*. The University of Chicago Press.
- Rahman, Z. A., & Agus, M. (2022). Kekerabatan bahasa Makassar dan bahasa Selayar: Analisis leksikostatistik dan glotonomi. *Gema Wiralodra*, 13(1), 215–232.
- Sulistyono, Y., & Fernandez, I. Y. (2015). Penerapan teknik leksikostatistik dalam studi komparatif bahasa Baranusa, Kedang, dan Lamaholot di Nusa Tenggara Timur. *Jurnal Penelitian Humaniora*, 16(1), 1–13. <https://doi.org/10.23917/humaniora.v16i1.1506>
- Tadmor, U., Haspelmath, M., & Taylor, B. (2012). Borrowability and the notion of basic vocabulary. *Diachronica*, 27, 226–246.
- Wiemer, B., Wälchli, B., & Hansen, B. (Eds.). (2012). *Grammatical replication and borrowability in language contact* (Trends in Linguistics: Studies and Monographs, Vol. 242). De Gruyter Mouton.