

Communicating the Keywords of Sendai Framework: A corpus-based approach

Sivadass Thiruchelvam^a, Ng Yu Jin^b, Chong Seng Tong^c, Norhayati Bte Mat Husin^d, Azrul Ghazali^e, Ahmad Kamal^f, ^{a,b,c,d,e,f}Universiti Tenaga Nasional, Selangor, 43000, Malaysia, Email: ^aSivadass@uniten.edu.my, ^bYuJin@uniten.edu.my, ^cStchong@uniten.edu.my, ^dHayati@uniten.edu.my, ^eAzrulg@uniten.edu.my

The Sendai Framework provides insights in terms of multiple levels of governance and recommendations on reduction of disaster risks. It is crucial that the essence of the framework is fully comprehended especially in communicating the priorities to various parties and stakeholders. The interpretation of the framework by various professional organisations, academia and stakeholders, for the purpose of planning and implementation, can be a challenging and laborious task. This paper will present an alternative perspective that simplifies the interpretation of knowledge dissemination (content analysis), especially when connecting the Sendai Framework key essence using a corpus-based approach. ‘Keywords’ will be used as a lens for community and stakeholders involved in disaster risk reduction planning or execution. This is achieved by a corpus-based investigation of the Sendai Framework Blueprint (official document). A keyword analysis was performed to identify words associated with the Sendai Framework core and of the potential pedagogic values therein. This paper suggests appropriate and continuous interpretation of the analysed content, in order to achieve improved disaster risk reduction communication within the community of stakeholders.

Key words: *Sendai framework, corpus-based studies, disaster risk reduction communication.*

Introduction

The Sendai Framework for Disaster Risk Reduction 2015-2030 (hereafter the Sendai Framework) was presented and passed at the Third United Nation (UN) World Conference on Disaster Risk Reduction on March 18, 2015. It was then endorsed by the UN General Assembly on June 3, 2015. The framework aims to reduce disaster risk and loss of lives, livelihood and health, as well as socioeconomic, socio-cultural and environmental assets concerning stakeholders (Toinpre, Gajendran & Mackee, 2018; Jaroslav, Swenja, Annegret, Reinhard & Jeroen, 2016). It is a fifteen--year, voluntary, non-binding agreement which emphasises the collective roles and shared responsibilities in reducing disaster risk, and includes the local government, the private sector and its stakeholders (Munenea, Swartling & Thomalla, 2018). It has four main and meaningful priorities, summarised as follows:

1. Understanding disaster risk and vulnerability, which includes knowledge relating to risk assessment, prevention, mitigation, preparedness and response.
2. Strengthening disaster risk governance to manage disaster risk at multilevel governance, and nationals, in a manner that fosters collaboration and partnership.
3. Public and private investing in disaster risk reduction for resilience, in order to promote socioeconomic resilience, as well as assets and environment protection.
4. Enhancing disaster preparedness for effective response, and for “Recover and Rebuild Better” in recovery, rehabilitation and reconstruction, which includes integrating disaster risk reduction into development measures.

In attempting to communicate the essence of this framework to various parties, another perspective of ensuring the right message is disseminated can be initiated using the content analysis method (corpus-based studies, cf. Paltala, Boano, Lund & Vos, 2012). This type of study identifies the ‘anatomic’ core words (lexis) in the Sendai Framework texts, or blueprint. In other words, the keywords of this framework can be ascertained and can contribute to the consolidation of knowledge or subject matter. The linguistic data is believed to demonstrate the ‘natural language’ of the framework which affects its implementation.

Literature Review

Hsieh and Shannon (2005, p.1278) define qualitative content analysis as “a research method for the subjective interpretation of the content of the text data through the systematic classification process of coding and identifying themes or patterns”. Qualitative content analysis has also been defined as “an approach of empirical and methodological controlled analysis of texts within their context of communication, following content analytic rules and step by step models, without rash quantification” (Mayring, 2000, p.2).

Content Analysis

According to Ary, Jacobs and Sorensen (2010), content analysis is “a research method applied to written or visual materials for the purpose of identified specified characteristics of the material” (p. 457). It can also be said that “content analysis entails a systematic reading of a body of texts, images, and symbolic matter, not necessary from an author or user’s perspective” (Krippendorf, 2004, p. 3). All in all, it can be inferred that content analysis is a systematic approach or technique that can be carried out in a manner where the results can be verified or replicated by other investigators who follow the suggested steps of the original researcher (Krippendorf, 2004).

Content analysis can also be quantitative in nature. Hardy, Harley and Philips (2004) define corpus-based approach analysis as objectively belonging to the quantitative analysis sphere, with qualitative characteristics. Similarly, Conrad (2005) advocates that analysis which involves identification of language patterns or trends necessitates a quantitative assessment to determine those patterns which are more frequent than unusual or unique cases of language use. Quantitative content analysis refers to “summarizing quantitative analysis of messages (in texts) that emphasize objectivity, systematicity, reliability, generalizability and replicability” (Neuendorf, 2004, p.33). It is systematic as there are clear rules based on theories or assumptions that include and eliminate certain content or diagnostic features (Zhang & Wildemuth, 2009). The analysis includes examination of texts, frequency counts and statistical measurements on texts which are abstracted from their contexts.

Corpus linguistics

In the field of corpus linguistics, a corpus can be defined as a collection or pool of words associated with the digitising of texts, which is often referred to as lexis (Nesi, 2013; Nation, 2001). In other words, the words present in a corpus must be converted into an electronic database, so that they are machine-readable (Tognini-Bonelli, 2001), as well as considered to be of “real-life language use” (McEnery & Wilson, 2001, p.1). “Corpus linguistics is the study of a language use on the basis of texts with source of evidence for linguistic description and argumentation” (Kennedy, 1998, p.7). With all these definitions in mind, in this study, the term corpus refers to the combination of words, phrases and the language used in the whole text. In other words, a corpus describes the properties and patterns of how a particular language is used. A corpus is a collection of texts which, upon examination of the ‘anatomy’ of the texts; provides more explicit properties of the language used holistically in context (Menon, 2009; Mudraya, 2006; Mukundan, 2009; Ward, 2009; Wang, Liang & Ge, 2008; Willis, 1998).

When compared with the unavailability of computers over the last three decades, the emergence of computers that produce computerised corpora has elevated the ‘scholarly

enterprise called corpus linguistics' (Fung, 2003). Flowerdew (2012, p.320) asserted that "this empirical approach investigates naturally occurring language, i.e. performance-based data". With the advance of computer technology, corpus linguists are able to analyse large diversities of texts and seek a broader view of language use. Computers facilitate empirical analyses of the actual complex patterns of language use, finalising diverse linguistic decisions (Biber et al., 1998). Therefore, the building of a corpus is indispensable in providing quantitative and empirical evidence of 'scientific' generalisation. The advantages of corpus-based studies include "reduced speculation and subjectivity; authenticity of data; and the potential to verify research hypotheses systematically and based on more extensive linguistic material" (Biel, 2009, p.2). In the latest development, word stemming is proposed to identify various morphological variants (Singh & Gupta, 2019).

Keywords Analysis

Scott (1999; 2001; 2004; 2008) applies the statistical measurement of the Log Likelihood Test to compute and identify key words in a corpus through the 'KeyWords' function. The Log Likelihood Test refers to a statistical assessment that compares two models to show how frequently certain data is likely to occur than the other. In other words, the model offers a better evaluation of the essential keywords used with the most likelihood of occurring in the corpus. The Log Likelihood Test also offers the best possible alternative test (Chi Square Test) to identify the key words of a target corpus, compared with a much larger corpus (Baker, 2010; Baron, Rayson & Archer, 2009; Rayson, Berridge & Francis, 2004; Rayson, 2003). The KeyWords function is used to identify or 'filter out' keywords in a given text which are 'unusually' high in frequency as compared to the norm (Bondi & Scott, 2010; Baker et al., 2008; Xiao & McEnery, 2005).

Any single word with 'normal high frequency' or is 'outstanding' in terms of its frequency in the texts, is considered as "key". The order of the key words is based on the order of 'keyness', which then produces a rank of high frequency words. There are two types of words with 'keyness' or keywords. The ones which are normally frequent in the target corpus when compared with the reference corpus are termed as positive keywords. In contrast, negative keywords are those that are infrequent in the target corpus in comparison with the reference corpus. Only words with positive 'keyness' are engaged as the keyness value signals the extent to which a specific word is more frequent.

Purpose of the study

The Sendai Framework provides insights in terms of multiple levels of governance and recommendation on reduction of disaster risks. It is crucial that the essence of the framework is fully comprehended, especially when communicating the priorities to various parties and

stakeholders. The interpretation of the framework by various professional organisations, academia and stakeholders – in relation to planning and implementation - can be a challenging and laborious task. This paper presents an alternative perspective that simplifies the interpretation of knowledge dissemination (content analysis), especially in connecting the Sendai Framework key essence using a corpus-based approach. This approach uses ‘keywords’ as a lens for community and stakeholders involved in disaster risk reduction planning or execution. In addition, using a corpus-based approach, the idea of eliminating vague language is possible (Li, 2019).

Due to the importance of the Sendai Framework, it is therefore meaningful to investigate the key lexical items used in creating the official manual, or blueprint, in order to deliver the intended ‘comprehension’ effectively. Specifically, the following research questions are addressed:

- (i) *What will be the lexical features that are significant in the official document of Sendai Framework?*
- (ii) *What are the essential keywords identified?*

Methodology

This study focuses on the linguistic aspect of content analysis, especially the analysis of running texts or lexical items. The methodological platform of this corpus-based research covers the fields of corpus linguistics and content analysis in the process of examining the lexical properties and the relationships between the lexical items in the text. Hence, this corpus-based study uses the quantitative content analysis approach, due to the nature of corpus linguistic studies (which deals mostly with numbers and percentages of use). This study comprises objective measurements, and statistical data analysis which deals with statistical measurements that are produced by corpus analysis or concordance software (concordancer). The statistical data includes frequency count, range figures, percentages, keyness (keyword identification) and lexical coverage. It is necessary to note that the analysis carried out is a corpus-based analysis, which is descriptive in nature and focuses on lexis. The analysis of the study aims to provide an in-depth insight through an extensive investigation of the lexical properties of a Sendai Framework document. To obtain the statistical data on frequency, percentages and keyness (using log-likelihood test statistics), the concordance software of WordSmith Tools 5.0 was used. Thus, the comparative frequency approach of using the reliable Log-Likelihood Tests found in the WordSmith Tools 5.0 was used to provide more insightful frequency data to determine the keywords (Scott, 2001; 2008; Scott & Tribble, 2006; Rayson, Berridge & Francis, 2004).

In order to perform this analysis, one should compare the words of target texts with those of a reference corpus that is taken from a much larger corpus. A reference corpus can be a benchmark or standard reference which has a corpus size ranging from a few thousand to hundreds of millions of words and is capable of representing the collection of a natural target language (Baker, 2006). It is affirmed that a reference corpus has been designed to provide universal information about a language (Leech, 2002; Sinclair, 1991). In other words, a reference corpus must be large and relevant enough to represent most, if not all, of the varieties of the target language and its characteristics, as well as being at least five times greater than a comparison corpus (McEnery, Xiao, & Tono 2006; Berber-Sardinha, 2002). The British National Corpus (BNC) which has been used as a reference corpus in several corpus-based related studies (see Ng et al., 2013; Mukundan & Ng, 2012; Al-Marooqi et al. 2011) was the ideal reference corpus selected for this study.

Results and discussions

Statistical distribution

The tokens (known as running words) and types of the target corpus (Sendai Framework) can be derived from the concordance software. Statistical distribution from the software provides detailed information of the loading of words and lexical density with its recurring ratio. From the concordance software the following characteristics could be identified: The number of tokens (also known as running words), the kinds of words (referred to as types), the type/token ratio (TTR) and the standardised TTR. The type/token ratio (TTR) is the number of types per token, giving an outline of the lexical density of words which is often labelled as density ratio (Ng, et al., 2013; Al-Marooqi et al., 2011; Mukundan, 2009). By identifying the types per tokens ratio, as defined by Laufer and Nation (1995), the lexical variation (LV) of the Sendai Framework corpus would provide a better outlook to the properties of the texts. Texts with very high lexically varied choices can debilitate the comprehension process (Mukundan, 2009; Tomlinson, 1998). In this respect Schmitt (2002) notes the importance of lexical variation in the diversity of the vocabulary used in a corpus. Hence, the higher the TTR, the more lexically varied the corpus is and the more difficult the texts can be for readers to comprehend. In a study, Kuo (1999) found that the TTR in science journals ranges from 4 to 7 based on the simplistic type-to-token ratio calculation. This study shows that the TTR in the Sendai Framework is even higher than those of the science journals found in the Kuo (1999) study.

Table 1: General Statistics obtained

| Corpus | Tokens | Types | Density [Types/Tokens Ratio (TTR)] | Ratio | Standardized [Types/Tokens Ratio (STTR)] |
|------------------|--------|-------|---|-------|---|
| Sendai Framework | 15897 | 1508 | 9.49 | | 33.85 |

On the other hand, it is interpreted that a high type-to-token ratio illustrates little repetition of each word in the corpus, meaning that it has a high density of words variation. Thus, it indicates that a particular text has a higher level of difficulty with a wider range of vocabulary. The TTR, as proposed by Schmitt (2002), is equivalent to that of the Standardised TTR (STTR) obtained from WordSmith 5.0, in terms of the dissimilar token size and the inconsistent token-to-type proportioning. In other words, the simplistic calculation of the TTR in this study is shown for reference purposes. In addition, the STTR was used as the meaningful data in this study to gauge the lexical density. If the STTR is valued at 34, for example, it illustrates that there are 34-word types in every 1000 words in the corpus. The higher the STTR would suggest the higher lexical variations or diversities in the target corpus (Scott, 2001; 2008). Noorli Khamis and Imran Ho Abdullah (2013) suggested that a low STTR would mean that many similar types of words found are largely repeated and the texts can be easily comprehended. According to Schmitt (2002), the appropriate ratio for written texts should range from 36 to 57. It can be summarised that the texts in the Sendai Framework may be somewhat overwhelming or confusing for readers for whom English is not their first language. Although there are translated versions of this text, the original base is still the English version.

Keywords identified

The keywords identified are words which are “unusually” high in frequency as compared to a reference corpus (BNC). These words are the core words which encompass the essence of the text of the Sendai Framework. There are 288 words identified which can be found in the appendix. Ten of the most frequently used (10 most positive keyness words) are discussed. They are:

“DISASTER, RISK, REDUCTION, RESILIENCE, STAKEHOLDERS, COOPERATION, FRAMEWORK, IMPLEMENTATION, GLOBAL (COUNTRIES), DEVELOPMENT”.

It can be observed that the keywords are parallel with the objectives, mission, vision and even goals of the Sendai Framework. It can be interpreted quite objectively that the framework targets “disaster risk reduction” and no matter how we look at these words, they are truly in line with the core beliefs of the framework, and the intention of improving lives throughout the world. By focusing on the keywords identified, stakeholders or “readers” of the blueprint could truly internalise the ‘core message’ of the framework (Palttala, Boano, Lund & Vos,



2012). The rest of the words identified are in many ways essential to communicating the Sendai Framework to a mass public and could perhaps be included in certain curriculums.

Limitations of the Study

The investigation of the study is subject to the corpus linguistics and lexical approach framework where the researchers will dictate the data solely based on the concordance software. There was no interview and questionnaire data involved, but this could be a future recommendation indicated by this research.

Conclusion

The relationship between the corpus linguistics, keywords and the texts of the Sendai Framework has been demonstrated as equating the language of communicating the framework with the keywords identified from its corpus. The results show the importance of knowing the keywords, in order to ensure that the right ‘implementation or action’ is eventually taken or emphasised in any plan. Keywords are closely related to the specific domain or discourse community as well, as there are a high percentage of key terms which are worth noting. The unusually high frequency contributes to consolidated action or communication as well as facilitating the understanding of the subject matter.

Acknowledgement

The authors are grateful to **Tenaga Nasional Berhad, Malaysia and Universiti Tenaga Nasional** for the funding of this project.

REFERENCES

- Al-Mahrooqi, R., Al – Busaidi, S., Mukundan, J., Ahour, T. & Ng, Y.J. (2011). Can the essential lexicon of geology be appropriately represented in an intuitively written EAP module? *English Language and Literature Studies*, 1(1), 50-66.
- Ary, D., Jacobs, L. C. & Sorensen, C. (2010). *Introduction to Research in Education* (8th ed.). California, USA: Wadsworth Cengage Learning.
- Baker, P. (2006) *Using corpora in discourse analysis*. London, UK: Continuum.
- Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh, UK: Edinburgh University Press.
- Baker, P., Gabrielatos, C., Khosravinik, M., Kryzanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourse of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273-306.
- Baron, A., P. Rayson, and D. Archer. (2009). Word frequency and key word statistics in corpus linguistics. *Anglistik: International Journal of English Studies*, 20(1), 41–67.
- Berber-Sardinha, T. (2002). Comparing corpora with WordSmith Tools: How large must the reference corpus be? Retrieved 13 November 2011 from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.7923>
- Biber, D., Conrad, S. and Reppen, R. (1998). *Corpus linguistics: investigating language structure and use*. Cambridge, UK: Cambridge University Press.
- Biel, Łucja (2009) *Corpus-based studies of legal language for translation purposes: Methodological and practical potential*. In H. Carmen & E. Jan (Eds.), *Reconceptualizing LSP*. Online proceedings of the XVII European LSP Symposium 2009. Aarhus 20.
- Bondi, M. & Scott, M. (2010). *Keyness in texts*. Amsterdam: Benjamins.
- Conrad, S. (2005). *Corpus linguistics and L2 teaching*. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp 393 – 409). Mahwah, NJ: Lawrence Erlbaum Associates.
- Flowerdew, L. (2012). *Corpora and language education*. Hampshire, UK: Palgrave Macmillan.
- Fung, L. P. Y. (2003). *The use and teaching of discourse markers in Hong Kong: students' production and teachers' perspectives* (Doctoral dissertation, University of Nottingham)
- Hardy, C., Harley, B. & Phillips, N. (2004). *Discourse Analysis and Content Analysis: Two Solitudes?* In Y. M. Herrera, & B. F. Braumoeller (Eds.) *Symposium: Discourse and Content Analysis*. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 2(1), 19–22.
- Hsieh, H.-F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Jaroslav, M. Swenja, S. , Annegret, T., Reinhard, M. & Jeroen, A. (2016). Brief communication: Sendai framework for disaster risk reduction – success or warning sign for Paris? *Natural Hazards and Earth System Sciences*, 16 (10). pp. 2189-2193.



- Kennedy, G. D. (1998). *An Introduction to Corpus Linguistics*. London, UK: Longman.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: SAGE.
- Kuo, C. H. (1999). Can numbers talk? Basic data management in a corpus. *RELC Journal*, 30 (1), 1-17.
- Laufer, B. and Nation, P. (1995). Lexical richness in L2 written production: Can it be measured? *Applied Linguistics*, 16 (3), 307-322.
- Li, S. (2019). Communicative significance of vague language: A diachronic corpus-based study of legislative texts. *English for Specific Purposes*, 53, 104-117.
- Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1(2). Retrieved July 28, 2014, from <http://217.160.35.246/fqs-texte/2-00/2-00mayring-e.pdf>.
- McEnery, T. & Wilson, A. (2001). *Corpus linguistics: An introduction* (2nd ed.). Edinburgh: Edinburgh University.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies. An advanced resource book*. London: Routledge.
- Menon, S. (2009). *Corpus-based analysis of lexical patterns in Malaysian secondary school Science & English for Science & Technology textbooks*. Unpublished doctoral dissertation, Serdang: Universiti Putra Malaysia, Malaysia.
- Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purpose*, 25, 235-256. doi:10.1016/j.esp.2005.05.002
- Mukundan, J. & Ng, Y.J. (2012). Development of a technical nursing education word list (NEWL). *International Journal of Innovation in English Language Teaching and Research*, 1(2), 105-124.
- Mukundan, J. (2009). *ESL textbook evaluation: A composite framework*. Köln, Germany: Lambert Academic Publishing.
- Munenea, M.B., Swartling, A.G. & Thomalla, F. (2018). Adaptive governance as a catalyst for transforming the relationship between development and disaster risk through the Sendai Framework? *International Journal of Disaster Risk Reduction* 28, 653–663
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nesi, H. (2013). ESP and corpus studies. In B. Paltridge & S. Starfield (Eds.). *The handbook of English for Specific Purposes* (pp. 407 – 426). Sussex, UK: John Wiley and Sons Inc.
- Neuendorf, K.A. (2004, Spring). Content analysis: A contrast and complement to discourse analysis. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 2, 1
- Ng, Y. J., Lee, Y.L., Chong, S.T. Nurhanis Sahiddan, Philip, A. Noor Hafiza Nor Azmi & Mohd Ariff Ahmad Tarmizi. (2013). Development of the engineering technology word list for vocational schools in Malaysia. *International Education Research*, 1(1), 43-59.
- Noorli Khamis & Imran Ho Abdullah (2013). Wordlists analysis: Specialised language categories. *Pertanika Journal of Social Sciences and Humanities*, 21 (4), 1563 – 1581.



- Palttala, P., Boano, C., Lund, R., & Vos, M. (2012). Communication gaps in disaster management: Perceptions by experts from governmental and non-governmental organizations. *Journal of Contingencies and Crisis Management*, 20 (1), 1-12. doi:10.1111/j.1468-5973.2011.00656.x
- Rayson, P. (2003). *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. A published PhD thesis. Lancaster: Lancaster University
- Rayson, P., Berridge, D. and Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. In *Proceedings of the 7th International Conference on Statistical analysis of textual data (JADT 2004)*, Louvain-laNeuve, Belgium. pp. 926-936.
- Schmitt, N. (2002). Using corpora to teach and assess vocabulary. In Tan, M. (Ed.), *Corpus studies in language education* (pp. 31-44). Thailand: IELE Press.
- Scott, M. & Tribble, C. (2006). *Textual Patterns: keyword and corpus analysis in language education*. Amsterdam: Benjamins.
- Scott, M. (1999). *WordSmith Tools version 3*. Oxford: Oxford University Press. ISBN 0-19-459289-8.
- Scott, M. (2001). Comparing corpora and identifying key words, collocations and frequency distributions through the WordSmith Tools suite of computer programs. In M. Ghadessy, A. Henry, & R.L. Roseberry, (Eds), *Small corpus studies and ELT* (pp. 47-67). Philadelphia: John Benjamins Publishing Co.
- Scott, M. (2004). *WordSmith Tools version 4*. Oxford: Oxford University Press. ISBN: 0-19-459400-9.
- Scott, M. (2008). *WordSmith Tools version 5*. Liverpool: Lexical Analysis Software.
- Sinclair, J.M. (1991). *Corpus, concordance and collocation*. Oxford, UK: Oxford University Press.
- Singh, J. & Gupta, V. (2019). A novel unsupervised corpus-based stemming technique using lexicon and corpus statistics. *Knowledge-Based Systems*. In press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work: Studies in corpus linguistics*, v. 6. Amsterdam: John Benjamins
- Toinpre, O., Gajendran, T. & Mackee, J. (2018). A conceptual frame using 'knowledge' as a lens for deconstructing the Sendai Framework Priority 1: Understanding disaster risks. *Procedia Engineering* 212, 181–189
- Tomlinson, B. (1998). *Materials development in language teaching*. Cambridge, UK: Cambridge University Press.
- Wang, J., Liang, S.L. & Ge, G.C. (2008). Establishment of a medical academic word list. *English for Specific Purposes*, 27, 442-458. doi:10.1016/j.esp.2008.05.003
- Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purpose*, 28, 170-182. doi:10.1016/j.esp.2009.04.001.



- Willis, J. (1998). Concordances in the classroom without a computer: Assembling & exploiting concordances of common words. In B.Tomlinson (Ed.), Materials development in language teaching (pp.44-67). Cambridge, UK: Cambridge University Press.
- Xiao, Z., & McEnery, A. (2005). Two approaches to genre analysis: Three genres in modern American English. *Journal of English Linguistics*, 33, 62-82.
- Zhang, Y., & Wildemuth, B. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), *Applications of social research methods to question in information and library science* (pp. 222-231). Westport, CT: Libraries Unlimited.
- Yu, C., & Schafer, B. W. (2007). Simulation of cold-formed steel beams in local and distortional buckling with applications to the direct strength method. *Journal of Constructional Steel Research*, 63(5), 581-590.
- Yu, W.W. and La Boube, R.A., (2010). *Cold-formed steel design*. John Wiley & Sons.

Appendix

Keywords of Sendai Framework

DISASTER
RISK
REDUCTION
RESILIENCE
STAKEHOLDERS
COOPERATION
FRAMEWORK
IMPLEMENTATION
GLOBAL
COUNTRIES
DEVELOPMENT
FINANCIAL
REGIONAL
MECHANISMS
SPECIFIC
NATIONS
PREPAREDNESS
SUSTAINABLE
HAZARD
NATIONAL
DEVELOPING
SCIENTIFIC
PROMOTE
HEALTH
RELEVANT



APPROPRIATE
COORDINATION
RECOVERY
INTERNATIONAL
INCLUDING
STRENGTHEN
POLICIES
LEVELS
RECONSTRUCTION
MANAGEMENT
STRENGTHENING
REDUCE
REHABILITATION
ENHANCE
CLIMATE
SUPPORT
ASSESSMENT
UNITED
MEASURES
COMMUNITIES
LOCAL
RESPONSE
PERSONS
GOVERNANCE
IMPACT
INSTITUTIONS
MULTI
SHARING
CAPACITIES
SECTORS
CAPACITY
RISKS
PUBLIC
ENVIRONMENTAL
HAZARDS
INFORMATION
OFFICE
FRAMEWORKS
PLATFORMS
PLANS



CHALLENGES
VULNERABILITY
ORGANIZATIONS
REDUCING
ASSETS
PRACTICES
PREVENTION
ACTION
ECONOMIC
SUBREGIONAL
ENHANCING
KNOWLEDGE
MULTISECTORAL
EFFECTIVE
INCLUSIVE
PROGRESS
ENGAGEMENT
TECHNOLOGY
STRUCTURAL
PRIORITY
POST
TECHNOLOGICAL
LIVELIHOODS
INFRASTRUCTURE
PROGRAMMES
EXPOSURE
DISSEMINATION
SOCIAL
ASSESSMENTS
SYSTEMS
STRATEGIES
COHERENCE
STATES
ASSISTANCE
DATA
THROUGH
RELIEF
CULTURAL
IMPLEMENT
RESPONSIBILITIES



PROTECTION
SENSITIVE
WARNING
DISAGGREGATED
COMMITMENTS
OUTCOME
IDENTIFIED
MODALITIES
TRANSBOUNDARY
LOSSES
BUILDING
INDICATORS
ACCESSIBLE
PERIODIC
LANDLOCKED
COPYRIGHTED
GEOSPATIAL
EFFICIENT
PRIORITIES
DISSEMINATE
FACTORS
SUBSTANTIALLY
EXISTING
COMMUNICATIONS
PRIVATE
AWARENESS
TRANSPARENCY
INSTRUMENTS
SCENARIOS
INTERGOVERNMENTAL
EDUCATIONAL
PRONE
MITIGATION
METHODOLOGIES
SETTLEMENTS
COLLABORATION
ENSURE
GENDER
ECOSYSTEMS
CRITICAL



STANDARDS
ADDRESSING
AFFECTED
BUILD
DEFNED
HYDROMETEOROLOGICAL
OFFICIALS
TOOLS
INNOVATION
FACILITIES
ASSEMBLY
RESOURCES
INTEGRATE
REGULATIONS
BILATERAL
PARTNERSHIP
MOBILIZATION
RESILIENT
TECHNICAL
TRANSFER
INSTITUTIONAL
MONITORING
CREATION
LAWS
INFORMED
ERADICATION
COMMUNITY
SECTOR
PRESENT
ACADEMIA
MORTALITY
LESSONS
FORUMS
INCENTIVES
SECTORAL
MULTILATERAL
COORDINATED
REVIEW
TERMINOLOGY
DRIVERS



PARLIAMENTARIANS
ADEQUATE
SAFETY
TIMELY
DISPLACEMENT
EMPOWERMENT
HOSPITALS
ACCESS
SUPPORTING
UNDERSTANDING
CHANGE
SERVICES
BASED
AUTHORITIES
MANDATES
ACHIEVE
GUIDANCE
REGISTRY
CONFERENCE
DISPROPORTIONATELY
DEVELOP
PLATFORM
RELATED
REGARD
EMPOWER
CONTRIBUTE
INTEGRATED
ENTITIES
ENGAGE
INVESTMENT
CHARACTERISTICS
ENCOURAGE
FOLLOW
ACCORDANCE
POVERTY
INVESTING
PERIODICALLY
APPROACHES
MIGRANTS
ENHANCEMENT



VARIABILITY
INFRASTRUCTURES
GOAL
GOVERNMENTS
SCALE
DEGRADATION
CONSULTATIONS
TAILORED
FACILITATE
TARGETS
GAPS
STOCKPILING
NEEDS
ROLES
ANNEX
LEARNED
UNDERLYING
UNIVERSALLY
DISPARITY
AVAILABILITY
GUIDING
ENHANCED
DEVELOPED
COASTLINES
ENVIRONMENTALLY
EDUCATION
PREVENTING
PRINCIPLES
PATENTED
CODES
CONTINUITY
MAPPING
UPDATE
ENVIRONMENT
RESEARCH
DISABILITIES
PARTNERSHIPS
LIVELIHOOD
DIMENSIONS
CAMPAIGNS



SOCIETY
COMPOUNDING
DRILLS
ENABLING
AREAS
NATURAL
FOOD
COMPLEMENT
SAFE
PREVENT
IMPLEMENTING
MORBIDITY
LOGISTICAL
CONTEXT
REQUIRES
PLANNING
PARTICIPATION
VOLUNTARY
SCIENCE
PROCESSES
INDIGENOUS
EFFORTS
STANDARDIZATION