

令和元年6月16日現在

機関番号：33901

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K02776

研究課題名(和文) A comprehensive analysis of the phonetic characteristics of English as a lingua franca

研究課題名(英文) A comprehensive analysis of the phonetic characteristics of English as a lingua franca

研究代表者

ギルナー リーア (Gilner, Leah)

愛知大学・法学部・教授

研究者番号：20558803

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：国際英語というドミナント・ボキャブラリー (DOVO)を見分けることを目的としコーパス分析に基づいて調査し、DOVOの規模と適用範囲の変動範囲が明らかになった。それから各DOVOは音声文字に表記し、母音体系が音韻論における要素の特徴に関して、機能的荷重 (FL)を勘案するの分析調査を行った。各母音体系のFLランキングを確立された。機能的荷重 (FL)を勘案する分析のおよび解釈的枠組みが、音韻論体系の探求にどのように活用できるのかを研究するものである。

研究成果の学術的意義や社会的意義

The data obtained from these analyses are of potential interest to researchers in various domains. This data adds a usage dimension to available descriptions of the English language. The data could also be used to further inquiry in issue related to speech processing and pedagogical models.

研究成果の概要(英文)：This research accomplished the following:

1. The dominant vocabularies (DOVO) in English was established by means of analysis of 11 spoken English corpora (8 World English varieties and 3 English as a lingua franca corpora) were identified. A range of variance in size and coverage of DOVOs was established by this data; 2. A series of functional load (FL) analyses were conducted on the transcribed forms of the DOVOs and FL rankings were produced for vowel systems of Canadian English, East African English, Hong Kong English, Jamaican English, Indian English, Irish English, Philippine English, and Singapore English; 3. Analyses indicate markedly uneven distributions among the sounds in the systems. Each system displayed idiosyncratic organizational tendencies. At the same time, certain trends were shared across systems such as relatively greater reliance on anterior-based articulations, generally in the upper quadrant of the vowel space.

研究分野：Applied linguistics

キーワード：functional load vowel systems world English varieties English phonology

1. 研究開始当初の背景

The English language is currently a global linguistic resource with more than 2 billion users around the world (Crystal, 2003). The broad and rapid spread of the language is often qualified as unprecedented by linguistics scholars (Albl-Mikasa, 2013; Crystal, 2003, 2004b; Hung, 2009; Matsuda, 2017; Seidlhofer, 2005). Galloway and Heath (2014, p. 386) observe that this state of affairs has "drastically changed the sociolinguistic landscape of English". The present investigation provides a corpus-based characterization of the English language primarily aimed at World Englishes (WE) phonology and English as a lingua franca (ELF) phonology. The functional load (FL) paradigm will be used to provide comprehensive usage-driven descriptions of some of the phonological systems at play in global English contact scenarios. The construct of FL provides a means by which to quantify the relative amount of work carried out by each element of a linguistic class (Hockett, 1955; King, 1967b; Mathesius, 1929). It has been used extensively in phonology, in particular, and has been shown to be an informative means by which to observe "...how much a language relies on phonological constructs" (Surendran & Niyogi, 2006: 11) as well as the nature and dynamics of the systemic relationships they form (Oh, Coupé, Marsico, & Pellegrino, 2015: 155). This investigation applies the FL paradigm to the study of linguistic varieties and produces data that is otherwise unavailable to the field.

2. 研究の目的

- (1) To provide a comprehensive phonetic characterization of global Englishes by means of a collection of deep analyses of state-of-the-art corpora representing world English (WE) varieties as well as English as lingua franca (ELF) interactional scenarios.
- (2) To develop an analytical framework and analysis tools that can be used to transform raw corpora data into an WE/ELF phonetic knowledge database. Importantly, these tools will be available to other researchers so that the database can be expanded collaboratively.

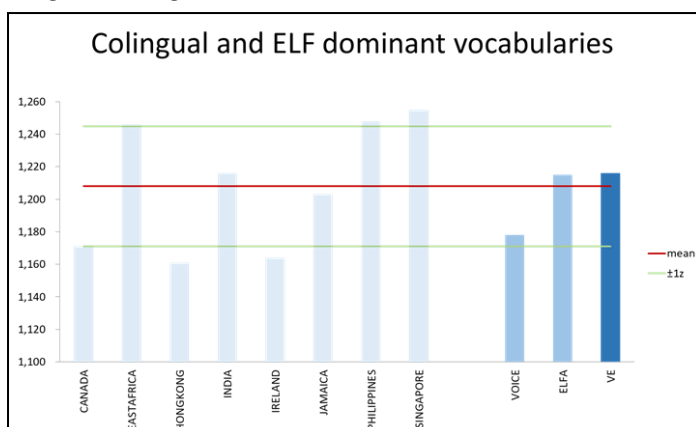
3. 研究の方法

- (1) The first year of this project entailed devising and implementing a methodology with which to extract the dominant vocabulary from electronic corpora. This methodology was used to analyze 11 spoken corpora, 8 representing as many varieties of world Englishes and 3 representing ELF. The 8 WE corpora were from the International Corpus of English (ICE): ICE-Canada, ICE-East Africa, ICE-Hong Kong, ICE-India, ICE-Ireland, ICE-Jamaica, ICE-Philippines, and ICE-Singapore. The ELF corpora were VOICE, ELFA, and the VE. The VE corpus was created by this researcher by combining VOICE and ELFA.
- (2) The second year was dedicated to creating a database of phonological forms for each DOVO. The absence of systemic and comprehensive data for the wide range of phonological systems at play in ELF interactions steered the investigation toward a population-based approach to analysis and description, resulting in a focus on WE varieties.
- (3) The third year was spent analyzing data and interpreting results. A series of analyses designed to identify, characterize, and classify the segmental components that comprise the ELF dominant vocabulary was devised and has proven informative. The methodology was applied to the analysis of several WE corpora, each corresponding to a different phonological system.

4. 研究成果

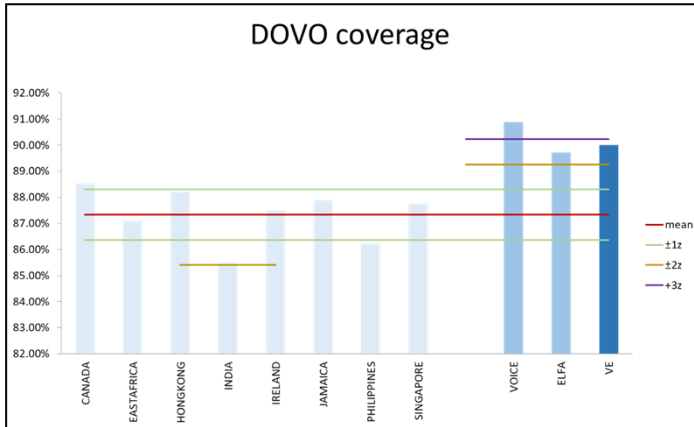
- (1) Lexical analyses of the ELF and WE corpora resulted in the identification of the vocabulary that dominate discourse in each corpus, referred to hereafter as DOVO. The WE corpora were used to establish a range of variance in size and coverage. Findings revealed that the dominant vocabulary identified in the ELF corpora fell slightly above this range of variance in terms of size. In other words, the ELF dominant vocabulary took up more of the lexical mass in its corresponding corpus than in the case of the WE dominant vocabularies. These results are displayed in Figure 1 and Figure 2.

Figure 1. Range of variance in DOVO size



As shown in Figure 1, the average colingual dominant vocabulary size is 1,208 words and the standard deviation is 36.84, plotted in the figure in the form of z-values. The green horizontal lines that cut longwise across the figure are 36.84 above and 36.84 below the 1,208 words on average. The coefficient of variation is 3.05%, a value that is generally considered to imply a narrow range of variation.

Figure 2. Range of variance in DOVO coverage



The coverage that each DOVO provides of its corresponding corpus is shown in Figure 2. With regard to colingual DOVOs, a minimum of 85.49% in the case of India to a maximum of 88.52% in the case of Canada is observed. A range of variation of approximately 3% can be observed among the colingual corpora. The average colingual DOVO coverage is 87.33%, the standard deviation is 0.96%, and the coefficient of variation is 1.1%. These results indicate that there is relatively less variation in DOVO coverage than in DOVO size among colingual corpora. Expressed differently, these metrics posit that DOVO coverage is more narrowly distributed than DOVO size.

The ELF corpora all fall outside the range of variation exhibited by colingual corpora and do so systematically above two z's. VOICE is 3.68 z's above the mean, closer in fact to 4 z's than 3. ELFA is 2.46 z's over the mean and VE is 2.77 z's away. To the extent that VE serves as a reference point for ELF, it is possible to conclude with reasonable certainty that ELF interactants rely more heavily on dominant vocabularies than colingual interactants.

(2) The second year of this research project focused on the design and implementation of a methodological approach appropriate to the description of phonological systems at play in ELF scenarios. An approach that provides the methods and means by which to make quantitative usage-based characterizations of each phonemic constituent in its various roles as one of many independent elements in a system, as a contrastive actor that establishes phonological oppositions with other elements, as a segmental member that participates in phonotactic sequences, and as a contributor to the amount of information conveyed by the system of elements was devised. A series of analyses designed to identify, qualify, and quantify the segmental components that comprise the ELF dominant vocabulary identified in first year of this research project were carried out on several phonological systems associated with a sampling of potential ELF users. The operationalization of functional load measures made it possible to observe idiosyncratic self-organizational tendencies in vowel systems. Furthermore, comparative analyses with control sets of low-frequency and random-frequency vocabulary sets revealed that dominant vocabularies distinguish themselves as phonologically highly-contrastive. A sampling of results are displayed in Table 1. Percentages indicate proportions in relation to the number of DOVO MPs. These findings contribute to the development and analysis of phonological models of English used as a global link language while augmenting current understanding of the nature of high frequency vocabulary.

		DOVO	Random		Low frequency	
Singapore English	Vowel MPs	1,116	560	50%	252	23%
	Consonant MPs	2,850	1,556	55%	820	29%
Jamaican English	Vowel MPs	816	444	54%	200	25%
	Consonant MPs	2,308	1,380	60%	788	34%

Table 1. Comparative assessment of MPs found in DOVO, random, and low-frequency words for two WE varieties.

(3) The third and final year of this investigation produced FL data for eight WE varieties. The research focused on the examination and characterization of vowel systems. Analyses targeted various levels of phonological organization including componential, systemic, and sub-syllabic. The outcomes of the research activities be summarized as follows: Phonological models for the eight world Englishes under investigation (viz. Canadian, East African, Hong Kong, Ireland, Indian, Jamaican, Philippine, and Singapore English) were validated based on a literature review relating to descriptions of phonological systems of world Englishes. Data sets comprising phonological forms of the dominant vocabulary of each variety were scrutinized and verified. A series of functional load analyses were carried out on each data set. An interpretative framework was devised in order to describe and discuss organizational patterns in terms of the usage intensity of articulatory gestures and phonetic features associated with the phonemes that constitute each vowel system under investigation. A platform to aggregate and query data sets was developed. An analytical tool that renders visual representations of results was developed.

A sampling of results is presented in Table 2. The results reveal a markedly uneven distribution of work across the 15 members of the HKE vowel system. The FL values range from a maximum of 0.0242 to a minimum of 0.0003. The member ranked 1st accounts for 19.8% of total amount of work carried out by this system. The 2nd ranked members

accounts for 13.18%, which means that these two phonemes carry out approximately 32.98% of the total. The members ranked 3rd to 5th contribute another 30.30% bringing the total for the top five members to 63.34%. The other ten members collectively contribute the remaining 37% to the total FL of this system. DRN values make it possible to describe FL in terms of fractions of the member with the highest FL. In this manner, the 2nd ranked member carries out 66.61% of the work the 1st ranked member does, that is to say, it contributes 66.61% of the work that the 1st ranked member does. These two members establish the upper- and intermediate-anterior perimeters of the vowel space. The 3rd, 4th, and 5th ranked members make lesser but rather similar contributions, each carrying out approximately half of the amount of work that the 1st ranked member does. The 3rd and 4th ranked members establish the lower- and upper-posterior perimeters of the vowel space. The 5th ranked member delineates the low-central perimeter.

Rank	Segment	MP types	FL	DRN	LDRN
1	i	14	0.0242	1.0000	85.55
2	eɪ	14	0.0161	0.6661	56.98
3	ɑ	14	0.0125	0.5145	44.02
4	u	14	0.0124	0.5118	43.79
5	aɪ	13	0.0122	0.5046	43.17
6	ɛ	13	0.0094	0.3885	33.23
7	ɔ	13	0.0062	0.2546	21.79
8	ɛə	14	0.0058	0.2409	20.61
9	oʊ	14	0.0051	0.2091	17.89
10	ʌ	12	0.0050	0.2057	17.60
11	aʊ	12	0.0041	0.1691	14.46
12	ɜ	14	0.0037	0.1542	13.19
13	ɪə	12	0.0030	0.1248	10.68
14	ʊə	9	0.0022	0.0898	7.68
15	ɔɪ	10	0.0003	0.0117	1.00

Table 2. FL ranking for the vowel Hong Kong English system

does a bit less than twice the work the least close counterpart. Among the posterior-based phonemes, however, there is very little difference between the FL values obtained by the most open member /a/ and most close member /u/, as just discussed. Closer inspection of the ranking shown in Table 2 indicates that anterior-based articulations are relatively most active. The high, front, tense /i/ contributes most to the work carried out by this system. The role of anterior-based articulations is further reinforced by the occupant of the next position in this ranking, the diphthong /eɪ/. Further scrutiny reveals that the seven members which populate the frontal region of the vowel space (including initial and terminal targets of diphthongs) collaborate to undertake 58.1% of the total amount of work carried out by this system. The 1st and 2nd ranked members account for more than half of that amount. The 15 vowel phonemes form MPs with 12 or 13 other members, on average 12.8. The least interactive member, ranked 14th, forms MPs with 9 others. Note that the maximum number of possible contrastive pairs or MP types is 14, there being 15 members in this system. Various members dispersed throughout the ranking are maximally interactive, that is, they form MPs with all other members, making it difficult to draw a direct relationship between FL and number of MP types. One might expect the more contrastively interactive members to have higher FL values but, as we will see when inspecting the results from the other varieties, the relationship between FL and number of MP types is not a simple one. This is so because FL values are ultimately derived from the frequency of occurrence of the word forms that participate in the MP under consideration for each contrastive pair. Simply stated, a given segment might form many but infrequent MP types, while another segment might form few but frequent MP types.

In summary, results of these analyses demonstrate how this FL can be applied to the domain of descriptive studies in order to provide a usage-driven perspective. Each of the systems investigated displayed distinct patterns of uneven distributions, indicating systemic relationships among members in terms of relative operability (Pellegrino, Marsico, & Coupé, 2011). These findings thus provide additional specific data that serves to systematically document the large diversity that is present among and within phonological systems. Several venues of potential application for this usage-driven data present themselves. When it comes to contemporary English studies, documentation and description are of particular relevance. The large body of referential work available on the language has traditionally been based on a limited and limiting perception of its users, uses, and usage. The aggregation of descriptive data sets can help advance the field in efforts to characterize the contemporary reality of the English language in its various forms and functions. Referential works which provide structural descriptions of phonological and grammatical features, encourage typological analyses and, as a consequence, afford a basis for more informed theorization regarding similarities and differences as well as speculating about the role of experience in cognitive categorization. Findings obtained from this research add to the resource pool available to researchers interested in pursuing such topics.

These findings are also relevant to substance-based theories of speech communication which position patterning of speech sounds in human languages as crucial to understanding productive/perceptual mechanisms underlying speaker-listener interactions (Liljencrants & Lindblom, 1972; Schwartz, Boë, Vallée, & Abry, 1997b; Stevens, 2002). The observed general tendencies in the world's languages of vowel systems toward symmetrical distribution of phonemes in the peripheral regions of the vowel space, for example, have been perceived as serving an economical energy function that satisfies both speaker demands for ease of articulation and listener demands of sufficient perceptual contrast

LDRN values express FL in terms of magnitudes relative to the member with the lowest FL. In this manner, it is possible to easily see that the 1st ranked member does 85.55 times the amount of work than the lowest-ranked member; the 2nd ranked member 56.89 times more; the 3rd ranked member 44.02 times more and the 4th and 5th ranked members approximately 43 times more work than the lowest-ranked one. These values expose a tendency toward close articulations. Among the anterior-based phonemes, the most close phoneme /i/ does 2.5 times the amount of work as the least close member /ɛ/; the intermediate member /eɪ/

(Schwartz et al., 1997b). The discussion of results made use of articulatory gestures to characterize the external manifestation of speech sounds by means of the physical movements underlying their production. Browman and Goldstein (1992: 23) identify gestures as dynamic articulatory structures "whose consequences can be observed in the movements of the speech articulators". Articulatory gestures embody the notion that speech sounds are complex structures associated with gestural patterns of speech, or production routines, that emerge through experience with usage events (Studdert-Kennedy & Whitney Goodell, 1992; Vihman, 2017). FL measures have been used to expose system-specific tendencies in gestural patterning among varieties of English. This data can be used to contribute to further development of substance-based theories of language.

Articulatory features have been used to address the internal representation that has been proposed by psycholinguists (e.g., Hayes, 2009; Hickok, 2014) and evidenced by brain studies (Chang, 2015; Eulitz & Lahiri, 2004). Usage-based phonology considers exposure and experience with situated and embodied meaning fundamental to understanding mental representations of phonological systems. A central concern is understanding and explaining how phonological categories are formed from highly variable speech tokens (Silverman, 2013). Contextual, physiological, and psychological factors conspire to produce "dramatic" acoustic invariance (Taylor, 2009, p. 23) that makes understanding how verbal communication succeeds "a major challenge" (Pitt, 2009, p. 19). Accumulated evidence demonstrates that a usage-based framework accommodates the invariance problem by allowing mental representations of phonological targets and patterns to gradually build up over time and with experience (Bybee, 1994; McQueen, Norris, & Cutler, 2006; McQueen et al., 2006; Norris, 2003; Pierrehumbert, 2001; Wedel, 2012b). The FL measures obtained from this investigation provide usage-driven quantification of hierarchies of phonological categories (and their associated physical attributes). These findings might thus be used to pursue inquiry into usage-based, exemplar-driven speech communication models that propose cognitive schemata emerge through "the collection of echoes" (Pierrehumbert, 2006, p. 523) retained from usage events.

Acknowledgment of the variation, variability, and variety that is inherent to language and communication is reorienting the linguistic sciences. When it comes to contemporary English studies, documentation and description are of particular relevance. The large body of referential work available on the language has traditionally been based on a limited and limiting perception of its users, uses, and usage. The adoption of a truncated view has had ramifications for subsequent theorization and speculation regarding relationships between language and the mind. Results discussed here encourage speculation regarding processing biases rooted in linguistic experience based on encountered exemplars, for which behavioral evidence continues to accumulate (Clopper, 2014; Clopper & Bradlow, 2007, 2009; Cutler, McQueen, Butterfield, & Norris, 2008; McQueen, 1991; McQueen & Cutler, 2010). Thus, it is tentatively proposed that FL analyses have something to contribute to discussions regarding language typology (Kortmann & Schneider, 2004; Schwartz, Boë, Vallée, & Abry, 1997; Schwartz, Moulin-Frier, & Oudeyer, 2015), language dynamics and evolution (Hruschka et al., 2009; Wedel, 2012b) as well as inclusive and pluricentric approaches to the study and teaching of the English language (Canagarajah, 2013, 2018; Cook, 2017; Kirkpatrick, 2010).

5. 主な発表論文等

[雑誌論文] (計 7 件)

- ① Gilner, Leah, A functional load quantification of the vowel systems of four Asian Englishes, (currently under revision for resubmission), 査読有, June 2019.
- ② Gilner, Leah, A usage-based approach to the description of vowel systems of world English varieties, (under review), 査読有, June 2019.
- ③ Gilner, Leah, A functional assessment of the Philippine English vowel system, 愛知大学言語教育研究室紀要「言語と文化」第 40 号、2018, 11-28.
- ④ Gilner, Leah, A glance into the mechanics of phonological signaling. 「アジア英語研究」, 査読有, 第 20 号、2018, 47-66.
- ⑤ Gilner, Leah, Language ecologies and world English varieties: Challenging implications, 文京学院大学紀要, 査読有, 第 17 号, 2018, 1-11.
- ⑥ Gilner, Leah, Intelligibility and global English-language use: An introduction, 文京学院大学紀要, 査読有, 第 16 号, 2017, 13-22.
- ⑦ Gilner, Leah, Identification of a dominant vocabulary in ELF interactions, *Journal of English as a Lingua Franca*, 査読有, 5(1), 27-51.

[学会発表] (計 5 件)

- ① Gilner, Leah, A functional assessment of vowel systems of several varieties of world Englishes, 24th Conference of the International Association for World Englishes, 2019.
- ② Gilner, Leah, Corpus-based descriptions of certain aspects of the lexical and phonological characteristics of eight varieties of World Englishes, The 23rd Conference of Pan-Pacific Association of Applied Linguistics, 2018.

- ③ Gilner, Leah, Towards the description of phonological systems of world Englishes by means of functional load analyses, 40th National Conference of the Japanese Association for Asian Englishes, Chukyo University, 2017.
- ④ Gilner, Leah, Functional load analyses in descriptions of ELF phonology, 10th Anniversary Conference of English as a Lingua Franca 'ELF and Changing English', University of Helsinki – Helsinki, Finland, 2017.
- ⑤ Gilner, Leah, Negotiation-driven convergence on a lexical scaffolding in ELF interactions, 9th Anniversary Conference of English as a Lingua Franca 'Framing English within multilingual policies and practices', Universitat de Lleida – Lleida, Spain, 2016.

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

Experimental Software for Corpus/Phonological Research

https://leo.aichi-u.ac.jp/~gilner/corpus-phonological_research.html

6. 研究組織

(1)研究分担者

研究分担者氏名 : n/a

(2)研究協力者

研究協力者氏名 : n/a

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。