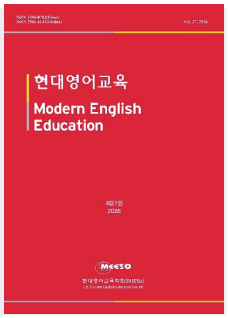
Modern English Education, Vol. 27, (2026) pp. 463-475
doi: https://doi.org/10.18095/meeso.2026.27.1.463

Modern English Education

ISSN: 1598-0782 / e-ISSN: 2586-6141
http://journal.meeso.or.kr



English Count Nouns: Variation and Statistical Tendencies

Kent A. Lee (Pukyong National University)

Received: 27 September 2026
Revised: 28 October 2026
Accepted: 16 November 2026

Lee, Kent A. (2026). English count nouns: Variation and statistical tendencies. *Modern English Education*, *27*, 463-475.

Keywords
Count nouns, grammar, noun phrases, definite articles, corpus
세가산명사, 문법, 명사구, 정관사, 말뭉치

Kent A. Lee
Associate Professor
Department of English Language and Literature
Pukyong National University
kentlee7@pknu.ac.kr
ISNI: 0000 0001 2758 354X

*This study was made possible by an Autonomous Creative Research Grant from Pukyong National University (2026), grant #202501111558.

Abstract

Standard approaches to teaching English nouns include traditional rules about count and non-count nouns, which do not necessarily reflect actual usage. While some nouns have tendencies toward count or mass noun usage, many do not, and many nouns show some degree of variability, depending on context and intended nuance. This project draws from an existing tagged corpus, using a Python script to classify nouns. From the resulting database of noun classifications, typical mass and countability preferences for nouns were calculated, leading to statistical indices for twenty thousand English nouns, and their likelihood of being used as count or mass nouns. These classifications can be useful for finding nouns that are more mass-like, flexible, or countable, for crafting teaching materials. The findings also demonstrate that English noun countability is better understood as a continuum rather than as a strict binary distinction, as many nouns exhibit flexible usage across count and mass contexts. Thus, traditional instructional approaches oversimplify actual noun usage patterns. Instead, instruction can benefit from showing typical noun phrase patterns, including those of more flexible nouns, and the changes in meaning or nuance involving count versus mass interpretations. A usage-based approach can contribute to a more realistic understanding of English noun usage and help learners to use noun phrases appropriately in context.

INTRODUCTION

In English grammar and writing instruction, the teaching of definite and indefinite noun phrases, and nouns with no article (bare nouns) remains a fairly under-explored and poorly understood area. English learners often encounter difficulty with English noun usage, e.g., whether a noun should be used in singular or plural form, and whether it can be used in count or mass noun phrases. In classroom instruction and in textbooks, these distinctions are frequently taught through categorical rules that classify nouns as count or non-count nouns, and through simplified lists of count and non-count nouns. However, a fair amount of research in corpus linguistics in past years has increasingly shown that many English nouns display considerable flexibility across contexts, appearing in different noun phrase patterns with different interpretations. For students, and for instructors of English as a foreign language (EFL), a more nuanced understanding of how nouns actually

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

behave in actual usage can yield better insights for explaining patterns of countability, definiteness, and number to learners. This study examines noun patterns in a large corpus to investigate how English nouns vary in their distribution across count and non-count noun phrase patterns. Frequency data on the use of nouns as count or non-count nouns in context can show to what degree nouns vary in their use as count or non-count nouns.

This study will attempt to quantify how nouns are used as count or non-count nouns, or both in varying contexts. By counting noun usage patterns in a large corpus, we can calculate a countability index for each noun. This can show which nouns are more strongly countable, which nouns are more strongly non-count, and which nouns can go either way, depending on the context and intended meaning or nuance. Countability index data for nouns can help educators and materials developers to better understand how various nouns are actually used. Based on this, teachers can then present noun patterns in a more realistic way that reflects how English speakers actually use nouns.

Thus, this study undertakes the following research question: How variable or flexible are English nouns between count and non-count usage? How common is this variability? Can this variability be quantified in a way that would be applicable for language educators?

LITERATURE REVIEW

English Nouns

Linguists have distinguished English mass (or non-count) and count (or countable) nouns since at least Jespersen (1924), and this distinction has been popularized in many descriptive and pedagogical grammar texts (e.g., Quirk et al., 1985). The main distinguishing criteria generally include the following: count nouns are pluralizable (e.g., *tree - trees*) while mass nouns are generally not so (*rice - rices*); count nouns take numeral modifiers (e.g., *two trees*), while mass nouns generally do not (**two honeys;* asterisks denote ungrammatical forms); singular count nouns take certain individuating quantifiers (e.g., *each, every, many, few*) while singular mass nouns instead take other quantifiers (e.g., *all, every, much, less*); count nouns can take indefinite determiners (*a cheetah*) and definite determiners (*the cheetah*), while mass nouns can only take definites (*the mud* cf. **a mud*); and finally, a singular mass noun alone can comprise a full noun phrase (NP or DP) (e.g., **Acorn / Rice spilled onto the ground*) (Fieder et al., 2014; Gillon, 1996; Jespersen, 1924; Krifka, 1995; Pelletier, 2012). Some tendencies have been noted, e.g., liquids and substances lack clearly defined physical boundaries and readily exist as mass nouns (*water, juice, plastic*), as well as small objects forming a collective or aggregate mass (e.g., *sand, rice*). Clearly defined objects readily exist as count nouns, as well as entities high on the animacy scale tend to be count (Krifka, 1995; Smith-Stark, 1974).

However, problems emerge with these generalizations, in that the mass-count noun distinction is not necessarily fixed (Krifka, 1995). Linguists have noted that the categorization is not necessarily fixed, as mass nouns can be coerced into use as count nouns for portions (*we ordered three coffees*) or taxonomic subtypes, i.e., for types or sorts of something (*a variety of cheeses, a fine wine, a strong linen*) (Krifka, 1995; Pelletier, 1975); conversely, count nouns can be coerced into mass nouns, e.g., in partitive expressions (*a lot of mango*) (Krifka, 1995; McCawley, 1975; Smith-Stark, 1974). However, some aggregates in English are count rather than mass nouns (e.g., *peas*), while some individuated entities are mass rather than count nouns (e.g., *broccoli, cauliflower*) (Vigliocco et al., 2002). Also, superordinate category nouns tend to behave as mass nouns, though they refer to objects or aggregates of objects (e.g., *furniture, jewelry*).

Flexible Nouns

Various linguists have noted that a number of nouns can exist as count or mass nouns, known as flexible nouns (Pelletier, 1975; Rijkhoff, 2008) or dual-life nouns (Payne & Huddleston, 2002). For example, Payne and Huddleston (2002, p. 335) famously noted common examples with object and substance nouns such as (*chocolate, a lot of chocolates, three chocolates*), *beer* (*drinking beer, ordering a beer*), and abstract terms (*much discussion, multiple discussions*). The flexibility of nouns is much more extensive than what some grammars and dictionaries (like learner dictionaries that categorize nouns as mass or count) let on, as many nouns can readily go between mass and count, and novel ad hoc semantic extensions can be created. For example, one small pragmatics study (Drożdż, 2020) examined nouns classified by English dictionaries as either mass or as count, and 1,700 native speaker utterances collected online; many nouns were found being used outside their typical taxa of mass or count, e.g.: (1) mass extensions of count nouns: "Dori naps next to the door. I think I got too much *door* in this photo" (p. 9), and "It's the eternal snack question: Why so much *bag* for so few chips, food companies?" (p. 10); (2)

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

count extensions of mass nouns: "A woman exercises on some aerial *silks*…" (p. 12); and (3) count extensions of aggregate (object mass) nouns, for a type or set of said aggregate: "Small offices and large offices will certainly find *a furniture* that will fit their specific office area" (p. 12), and "We pride our ability and willingness in providing high quality branded *equipments* at an honest and competitive price to all our customers, big and small" (p. 12).

Some psycholinguists have proposed that count nouns are encoded and processed as semantically individuated entities (Barner & Snedeker, 2005), and as a result, are semantically richer, more concrete, and more easily categorized. On the other hand, the interpretation of count-mass flexible nouns depends on context, reference, and morphosyntactic cues in context. Such hypotheses are supported by a number of language processing and production studies with children and adult English speakers, but oddities like aggregate object mass nouns and others above do not support a one-to-one mapping of objects with count nouns, or of substances with mass nouns (Barner & Snedeker, 2005, 2006; Fieder et al., 2014) and citations therein. Within a cognitive and philosophical framework, Pelletier (2012) proposes that nouns are not syntactically encoded as mass or count, but in their lexical semantics, can be either, depending on their usage in context; this attempts to account for the variety of flexible nouns and their usage in English. In contrast to the cognitive hypothesis that count nouns denote entities that speakers conceptualize as individuated and mass nouns for conceptually non-individuate items, others argue that the distinction depends on the distinctions in the speakers' minds, i.e., communicative function over inherent conceptual distinctions (Wisniewski et al., 2003). On the other hand, production and comprehension studies show evidence that English speakers use syntactic representations and cues like quantifiers and determiners to assess countability of nouns (Chiarelli et al., 2011; Fieder et al., 2014, 2018).

Learning and Teaching Issues

Some second language (L2) studies using grammaticality judgment tasks with Koreans learning English show that they are more sensitive to errors with count nouns (D. Kim & Kim, 2026), and more often overuse plural forms with mass atomic nouns (e.g., *furniture*) than mass substance nouns (e.g., water) (Choi et al., 2018). Such results suggest that atomicity or individuated object semantics is a default preference in a second language, or a language universal (Choi et al., 2018; Choi & Ionin, 2021). In a simple grammaticality judgment task, Koreans are better able to identify count nouns occurring in context than in isolation (S. Cho, 2006). In a study with Koreans performing a noun classification task (MacDonald & Carroll, 2018), Koreans were asked to classify English nouns as count nouns (e.g., *bees*), substance mass nouns (e.g., *lotion*), object mass nouns (e.g., *cattle*), or flexible nouns (e.g., *cake*, *cakes*). When identifying nouns in spoken sentence contexts with singular or plural noun forms, the Koreans had little difficulty with count and mass nouns, but compared to English speakers, had more difficulty with flexible nouns, particularly with their plural forms. The researchers attributed it to less familiarity with plural morphology; however, this could depend more on how they were taught English and a rigid mass versus count noun distinction, and are unaware of the flexible noun usages. Similar findings have been reported for Japanese learners of English for similar noun identification tasks (Inagaki, 2014), and flexible nouns are similarly problematic for Mandarin Chinese learners of English, based on an analysis of written essays (Antes, 2019). In grammaticality judgment tasks, Korean adults tended to prefer count interpretations of flexible nouns (Ma & Oh, 2014), and Koreans and Japanese are less likely to extend mass nouns to countable meanings (e.g., information, knowledge) than native English speakers (Min, 2014). In a forced choice task where Koreans had to categorize different types of abstract nouns as indefinite or Ø nouns (i.e., bare nouns, with no articles), participants were most accurate with bounded independent nouns (e.g., *fact*, *idea*) and least accurate with non-continuous action nouns (e.g., *jump*, *arrival*) (Amuzie & Spinner, 2013).

The difficulties of Koreans with English articles and nouns arise not only from language differences, but also due to classroom instruction. Pedagogy relies on simplistic mass-count noun distinctions and prescriptive rules on when one should or should not use definite and indefinite articles (H. Lee, 1999; K. Lee, 2021a). Flexible nouns and other complexities are generally not covered, and students report that the rules they learn seem not to work, or do not seem intuitively applicable (K. Lee, 2021b). Similar problems, such as not teaching flexible nouns, are common in China (Antes, 2019) and elsewhere.

While some learner dictionaries denote count (C) and uncountable (U) nouns, and a few may even denote some commonly flexible nouns as "C" and "U," for learners this information is inadequate, and for flexible nouns, it is not clear when the count and non-count uses are appropriate (Xue, 2010); this observation made in the context of Mandarin Chinese learners is valid for Korean, Japanese, and other learners whose L1 lacks an English-like article system. In fact, a study of Korean college students performing a diagnostic test on articles and think-aloud explanations afterwards shows confusion with flexible nouns and nouns varying from their typical count and mass noun uses (e.g., *an / the / Ø improvement*), which is partly due to inadequate and confusing dictionary information about count and non-count nouns (S. Kim, 2018).

Empirical data show that this categorical distinction is less than sound. One study (Grimm & Wahlang, 2021) using

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

machine learning methods for categorizing count and mass nouns from an early version of the COCA corpus (Davies, 2009) compared different criteria for identifying and classifying noun tokens, such as the presence of various premodifiers, quantifiers and determiners. Outcomes were compared with count/non-count classifications from the CELEX database (Baayen et al., 1995). It was found that bare singular nouns (mass) and bare plural nouns (count) were the most accurate and reliable criterion for classifying nouns from a corpus. However, the number of nouns was not indicated, the CELEX classifications are taken from Oxford and Longman dictionaries, and the noun classification dataset (Grimm & Wahlang, 2021) is not publicly available. Other studies have also explored machine learning algorithms for classifying nouns (Baldwin & Bond, 2003a, 2003b). Nouns have been found to be more elastic than generally assumed, as a machine learning classification study reports for count nouns (Katz & Zamparelli, 2012), and from a corpus-based study for abstract nouns (Zamparelli, 2020). The Bochum English Countability Lexicon, or BECL, (Kiss et al., 2016) classified 11,869 nouns into four semantic categories: count, mass, dual (i.e., flexible, or existing as both count and mass), and other (neither mass nor count, and study's criteria for this are not clear). The nouns were classified by raters based on grammaticality judgment tests and definitions from the WordNet database (Miller et al., 1990). Examination of the published dataset shows that 4,285 nouns had unique ratings, while the other 7,584 nouns were given multiple semantic ratings. While these ratings and dataset show the semantic variability or polysemy of nouns, they do not provide the type of data or detail that would be useful for pedagogy or other applications.

These corpus and lexicographically based comparisons have not led to publicly available noun type indices that could be useful for language teachers, materials developers, or applied linguists. Systematic comparison of mass, count and flexible nouns, and their degree of elasticity, seems lacking, especially for language teachers. The degree of noun flexibility across various nouns is also less clear, and at this point, language teachers or materials developers have limited guidance for teaching noun phrase patterns consistent with how native speakers use such nouns. Hence, this study attempts to create countability indices to facilitate the creation of appropriate materials and the selection of suitable nouns for instructional design.

METHOD

Data

Noun data are extracted from the Corpus of Contemporary American English (COCA), which consists of over one billion word tokens (1,001,610,938), from over 25 million words per year (1990-2019) from eight genres: fiction, popular magazines, newspapers, academic texts, blogs and other web pages, TV and film subtitles, and transcribed spoken samples (Davies, 2009, 2013). Each token is tagged with a part-of-speech (POS) or word class tags based on the CLAWS system used for other corpora such as the British National Corpus (Leech et al., 1994). This includes tagging nouns as singular or plural, or as proper nouns. The COCA corpus data comes in multiple forms, as multiple separate files, and as a single verticalized text file (VRT); the VRT file was used for this study, as it contains all the tokens, lemmas and tags in one file for straightforward data extraction.

Data Collection and Analysis

A number of tagging errors are known to occur in COCA, so several hundred errors were identified and corrected via regular expression search; this is especially so in the spoken corpora, due to the difficult nature of coding spoken data (Galiano & Semeraro, 2023)[1]. To expedite the process of data extraction, a Python script (Python version 3.14.2, running in Arch Linux version 6.18) was iteratively drafted and tested (with assistance from ChatGPT-4o), until a final Python script was settled upon. The code was reviewed by the researcher for accuracy, before being run to extract noun data from the corpus file. The script used the Python re (regular expressions) library, the csv library (for CSV files), and the sys library (for system-specific parameters and functions). The script extracted noun tokens, lemmas and tags, enumerated the sentences and tokens, and output the data into CSV files. Since random words in COCA are obfuscated for watermarking purposes, the script excluded such sentences. In the process, the script conducted the following checks and classified noun tokens in the output files.

Noun number was noted based on POS (part of speech) tags for singular or plural nouns. Compound nouns (based on adjacent POS tags for nouns) were noted as compounds, and were classified as non-head (non-final noun) or head noun (final noun, syntactic head of the NP). Proper nouns (based on POS tags) were marked for later exclusion. Quantifiers and

[1] See also https://userpage.fu-berlin.de/~structeng/wiki/doku.php?id=corpora:tagset-claws7-coxa for well attested typographical errors in COCA tags.

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

determiners before nouns were identified, to aid in classifying count and mass nouns. This involved using a lookback window, or looking up to six positions before the noun itself, to look for premodifiers and modifiers (adjectives, adverbs modifying noun phrase adjectives). The lookback stopped if it encountered a syntactic boundary, such as a verb, preposition, complementizer, or clause-level punctuation. Six positions were chosen based on examination of longer phrases in the corpus, to allow for phrases with stacked multiple adjectives and premodifiers. Within that window, premodifiers (articles, quantifiers, numerals, genitive markers, etc.) were noted. The lookback window allowed for commas and coordinating conjunctions within noun sequences (based on POS tags).

Based on these criteria and POS tags, general status was assigned (definite, indefinite, bare singular, bare plural, or indeterminate, i.e., could not be determined). Based on definiteness and number (from POS tags), countability was then assigned (count, mass, indeterminate). Nouns were classified as count nouns if they occurred in plural forms, or when modified by numerals or quantifiers associated with plurals (e.g., *many, few, several*). Nouns were classified as mass nouns when appearing as bare singulars or with singular modifiers (e.g., *much, little, some*). Nouns occurring less than five times in the corpus were excluded. Proper nouns (as indicated by POS tags) were also identified but excluded from subsequent analysis. When compound nouns were encountered, only the final noun (which is generally the semantic head noun) was chosen for exporting to the data file. The Python script wrote out data thus extracted into CSV files with token data and its classification of noun tokens. The script also attempted to further classify noun compound characteristics (for a separate research study), but generally did not do so accurately.

After removing watermarked sentences and tagging errors in the corpus, the usable COCA dataset contained 6.14 billion words and over 13 million sentences. The Python script found about 530 million possible noun tokens, after excluding about 2.2 million compound non-head tokens, 15.7 million proper nouns, and 315,800 noun tokens that could not be classified. The CSV output files were loaded into a SQL database (MariaDB 12.2 for Linux), along with noun concreteness ratings for 40,000 lemmas (Brysbaert et al., 2014). The Brysbaert ratings also contain general lexical frequency ratings (log frequencies) for lemmas, and standard deviations of ratings for concreteness, which can serve as a proxy for semantic variability. From the CSV files, nouns were extracted along with classifications for each token. A few erroneous rows had to be manually deleted, such as a few words that were not common nouns, whose inclusion apparently resulted from COCA tagging errors. This led to a set of 20,302 unique nouns. From these data, two noun countability indices were computed. For bare nouns, the proportion of bare plural nouns to overall bare nouns yielded a Bare Noun Countability Index (BNCI). The Python script's classifications of nouns as count or non-count from other criteria, for bare nouns and nouns in more complex noun phrases, yielded counts for uses of a noun as a count or non-count noun, and the proportion of count uses over all tokens of the noun yields a more comprehensive General Noun Countability Index (GNCI). The GNCI includes a larger set of nouns, and included most nouns in the BNCI. The output was then loaded into R (version 4.5.2) for brief correlational analysis (via the R psych library). Percentile rankings for the BNCI and GNCI were also calculated. The noun indices and rankings, noun frequency ratings based on the corpus, and the Python and SQL scripts, are available from the researcher's GitHub repository as a CSV file at https://www.github.com/kentlee7/nouns.

RESULTS

Noun Indices

In the script output, 20,302 unique nouns were classified as count or non-count based on various structural criteria, such as the presence of determiners, numerals, quantifiers, or such. Dividing count tokens over all tokens thus classified yielded the General Noun Countability Index (GNCI), a scalar index ranging from 0.0 to 1.0. A separate tabulation of simple noun phrases consisting of bare singular or bare plural nouns found a smaller set of unique nouns (12,693) that could be classified as count (bare plural) versus non-count usage, with the proportion of bare plurals over all bare nouns yielding a Bare Noun Countability Index (BNCI) with range 0.0-1.0. Since the GNCI includes most of the nouns, it serves as a more comprehensive index.

The BNCI and GNCI correlated with each other (see Table 1), and with data from Brysbaert et al. (2014), namely, lexical frequency, semantic concreteness (how English raters perceive a noun as tangible, physical, or mentally imageable), and semantic variability. Unsurprisingly, nouns rated as more concrete tend strongly to be higher in the countability indices from this study. The Brysbaert ratings should be considered cautiously, as they do not distinguish between very different meanings and lexical classes (parts of speech of words); for example, *maybe* occasionally appears as a noun in the COCA dataset (an instance of an indefinite response, e.g., "too many maybe's"), but its lexical frequency and concreteness ratings are skewed by its much more frequent adverbial usage. Nonetheless, simple Pearson correlations are shown below.

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

TABLE 1
Correlation Matrix

Factors	Semantic variability	Lexical frequency	GNCI	BNCI
Concreteness	-0.46	0.18	0.38	0.24
Semantic variability		-0.08	-0.10	-0.07
Lexical frequency			0.05	-0.15
GNCI				0.81

Note. All correlations were significant at $p < .0001$.

Since the classification procedure relied on the corpus part-of-speech annotations, the procedure was not able to perform more precise word-sense disambiguation. The indices will include occasional instances of lexical conversion or part-of-speech crossover (e.g., *maybe* used as a noun in *a definite maybe*). The indices should thus be understood as measures of lemma-level usage distributions rather than sense-specific measures of countability. Since these semantic conversions are likely low in frequency, users of the indices can open the CSV files in a spreadsheet or CSV viewer (e.g., Gnumeric), sort by particular indices, index values, and/or deciles, and then also sort by lexical frequency to see higher-frequency lexemes. Users who wish to choose very low-frequency lemmas may need to use common-sense judgments in choosing suitable nouns.

Data Validation

For some validation by comparison with other research data, the seemingly best available dataset is the Bochum English Countability Lexicon, or BECL (Kiss et al., 2016). The BECL was derived by raters rating nouns on their primary definitions and subsequent definitions in the WordNet database (Miller et al., 1990), and grammaticality judgments, into the following ratings or categories: count, mass, dual (i.e., flexible as count or mass), and an unclearly defined category of neither mass nor count ("other" in the tables below). While 4285 of the 11,870 nouns had unique ratings, the other 7584 nouns were assigned to two to four categories, with the first listed ranking per noun corresponding to a noun's primary WordNet definition. The BECL ratings were correlated with the decile ranks from the BNCI and GNCI indices via a chi-squared analysis (as the BECL ratings are nominal categorical variables, and the deciles are ranked categorical variables). The BECL categories correlated with the BNCI deciles ($\chi^2 = 190$, $df = 63$, $p < 0.0005$) and the GNCI deciles ($\chi^2 = 223.8$, $df = 63$, $p = 0.0005$). The general trend can be better seen in a cross-tabulation, e.g., between BECL categories and GNCI categories in Table 2; the BECL categories are represented as (only) count, primarily count plus other categories, and so on. The BECL count-only category tends to align with the higher GNCI deciles for higher countability, the category for count and others tends to correspond more with the medium GNCI deciles, and so on. The correlations are not exact; for example, although the BECL's count-only category generally tended to correspond to higher GNCI deciles, 1045 count-only nouns occurred in the lower four GNCI deciles. This is likely because the BECL raters only took the first four WordNet definitions for evaluation, so the BECL ratings may not reflect overall usage of polysemous nouns. While the BECL classifications draw on a limited number of lexical senses and grammaticality judgments, the GNCI reflects an empirically based distribution of count and mass uses in a corpus. These lower GNCI values may therefore reflect differences in actual sense distribution, actual contextual usage, and the differing classification procedures used for the BECL and GNCI.

TABLE 2
BECL Categories and GNCI deciles

BECL	GNCI deciles									
categories	0	1	2	3	4	5	6	7	8	9
Count (only)	213	285	285	262	302	383	399	464	433	311
Count + others	39	76	104	117	137	177	192	195	176	66
Dual (flexible)	1	4	6	11	7	7	2	2	2	0
Dual + others	5	23	32	38	47	45	53	56	37	24
Mass (only)	6	16	22	21	20	10	15	9	5	6
Mass + others	16	56	74	71	92	85	114	105	80	55
None (not mass or count)	1	3	2	2	3	2	1	1	0	1
None + others	3	10	18	31	28	26	36	23	32	6

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

The BNCI and GNCI seem accurate enough for educational applications. The upper deciles correspond to nouns that are more countable, the lower deciles to nouns that are more mass-like, and the middle deciles correspond to nouns that are more flexible. Thus, the nouns in the middle deciles show greater semantic variability and polysemy. For example, *glass* is in the GNCI 4th decile, as it can easily refer to a material or to an object (a drinking glass).

To evaluate the reliability of the Python-based noun classification procedure, a random sample of 400 noun tokens was drawn from the SQL database, for nouns used in computing the countability indices. The tokens were randomly sampled via a SQL query, and the unique token IDs were retained in the SQL output. Each sampled token was independently annotated by the author (who is a native English speaker) for contextual countability, number, and definiteness, following standard linguistic criteria, and without access to the corresponding automated classifications. The manual annotations were subsequently matched to the count classifications made by the Python script and their original sentences, by using the unique token identifiers for another SQL query. Agreement between the manual and automated countability classifications was evaluated using overall percentage agreement and Cohen's kappa (via the R irr library). Agreement between raters (the author and the Python script) for count and mass nouns was 93.2%, with Cohen's kappa=.79, indicating fairly reliable inter-rater reliability and a reasonable accuracy of the script in classifying nouns. Classification disagreements for 18 nouns stemmed primarily from three nouns that the researcher could not classify due to insufficient sentence output from the SQL query, nine nouns that the script could not classify, one compound noun tagging error in the corpus, and five flexible nouns, for which the script could not accurately classify the intended nuance in the context based on tagging. These five nouns (*effort, healing, service, society, surprise*) can be used for general states, general activities, and more abstract concepts, or for more specific types, events, or instances thereof, and understandably, a script cannot distinguish the particular nuance in context based on syntax alone. Otherwise, the script classification seems generally reliable.

Noun Distributions

Histograms and line area graphs (via the R psych, tidyverse and ggplot2 libraries) show the distributions of countability ratios, starting with the GNCI ratios below in Figure 1. The box histogram for GNCI ratios shows the number of nouns ranging from 0.0 (strongly non-count) to 1.0 (strongly countable) in 0.05 increments. A preponderance of nouns tends to be very or exclusively countable, as seen in the large counts toward the right. However, most of the rest of the nouns fall well in between the strongly count and strongly non-count ends. No discrete count / non-count categories are apparent; rather, nouns form a continuum or a spectrum ranging from count to non-count.

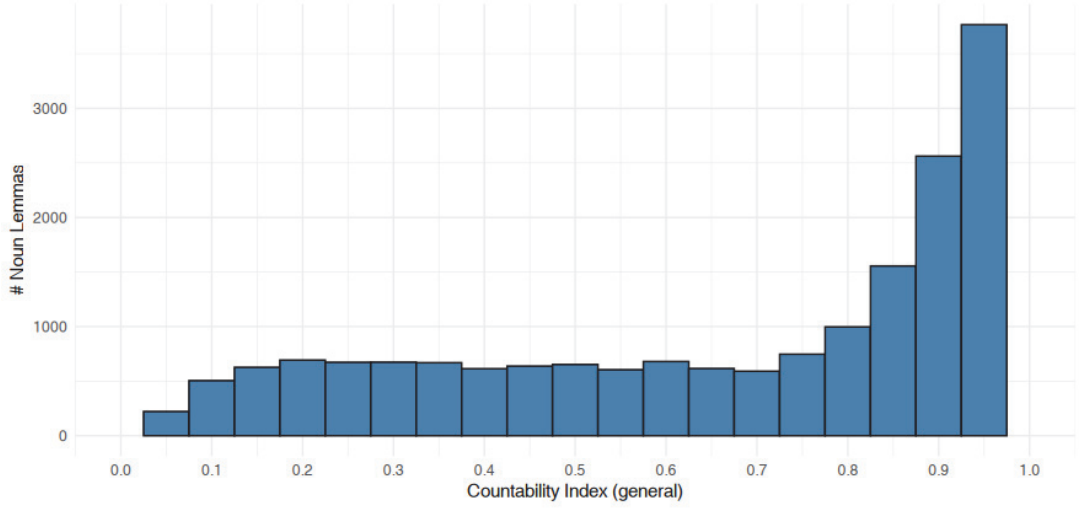


Figure 1
Distribution of General Noun Countability Index

The BNCI based on bare nouns misses 7589 nouns that appear in the GNCI (but includes 10 nouns that do not appear in the GNCI). Since it is based on a more binary classification of bare nouns, this shows a more bimodal distribution between both extremes, but also a more even distribution in between the two extremes of strictly non-count and strictly countable (see Figure 2).

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

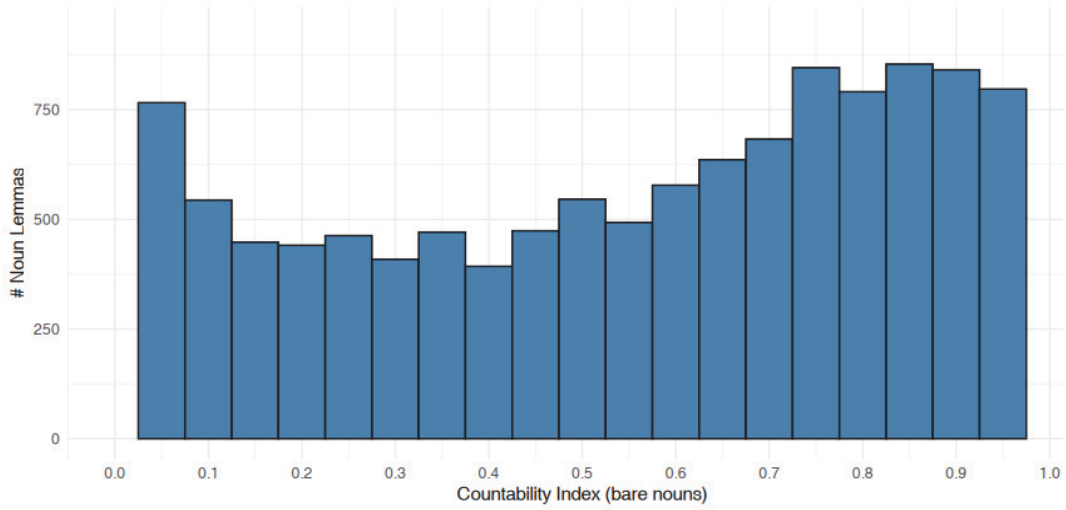


Figure 2
Distribution of Bare Noun Countability Index

Flexible nouns thus seem fairly common in English, as many nouns fall into intermediate levels with some degree of flexible usage. Some of the most frequently occurring nouns from the corpus are shown in Table 3, with the twenty of the most frequent nouns from each GNCI decile. The COCA corpus seems to capture more professional and academic vocabulary, as is especially apparent here.

Table 3
Higher Frequency Noun Examples Along Countability Continuum

GNCI decile	Top twenty words per decile
0	access, alcohol, baseball, basketball, care, discrimination, education, golf, honey, management, poverty, psychology, terrorism
1	cancer, congress, control, democracy, equipment, growth, health, information, justice, love, music, peace, politics, progress, research, security, television, training, trouble, violence
2	art, attention, college, color, county, death, development, director, earth, evidence, food, history, money, news, oil, photo, power, science, technology, water
3	action, age, air, culture, fact, figure, hair, home, interest, life, part, policy, practice, school, society, table, town, university, use, work
4	activity, business, center, change, city, class, court, experience, law, mind, morning, night, office, place, president, sense, service, time, view, war
5	baby, back, choice, church, department, effect, film, force, form, government, ground, heart, industry, kind, movement, officer, response, state, tax, type
6	boy, community, effort, face, family, game, girl, hand, head, level, line, mother, number, person, rate, reason, room, season, side, voice
7	body, car, company, difference, father, foot, group, house, leader, model, name, party, point, program, question, relationship, role, system, word
8	area, book, case, country, door, end, guy, hour, idea, issue, job, member, nation, player, problem, process, story, teacher, team, week
9	century, couple, decade, eye, friend, kid, minute, moment, month, official, parent, participant, patient, people, percent, police, resource, result, worker, year

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

Deciles 0-2 consist of nouns used mostly in a mass or abstract sentence, especially in Decile 0; these are almost always bare singular (i.e., with no article) abstract or material nouns, plus some nouns for general activities (e.g., *baseball*). Flexible usage increases with higher deciles; e.g., water in Decile 2 is usually mass, but occasionally count (*a water* for a serving or a bottle of water). Deciles 3-5 represent flexible nouns, which are more readily used in a count or non-count sense (e.g., *time, a time; a hand, by hand*), while the higher deciles increasingly represent nouns that are mostly countable nouns, e.g., especially Decile 9.

Aside from compound heads, non-head compound nouns were excluded from these indices, and treated separately. Nouns were classified by how often they formed compounds, and a ratio of compound tokens over all tokens of a noun lemma (compound and non-compound forms) was calculated to yield a noun compound frequency index (available from the researcher's GitHub repository). The most commonly compounded nouns in COCA include the following: *school, state, john, health, president, world, family, north, south, time, security, college, white, business, city, government, student, education, research, community, food, water, music, news,* and *drug.* However, the compound noun dataset (especially for non-head nouns) shows numerous errors and inconsistencies, due to tagging errors in the COCA corpus, so no further analysis was conducted. In the future, more research is needed on compound noun phrases.

DISCUSSION

The results of this study show that a number of nouns fit within intermediate ranges between mass and count, and these ranges allow for these nouns to be used as flexible, polysemous nouns that can take on mass or count interpretations according to context and speaker intent. This accords with the various qualitative and theoretical studies that argue for noun flexibility, and evidence therein from natural data and native speaker intuitions (e.g., Payne & Huddleston, 2002; Krifka, 1995; McCawley, 1975; Pelletier, 1975; Rijkhoff, 2008). This is also consistent with pragmatics research with natural data and native speaker judgments (e.g., Drożdż, 2020), and pragmatics research showing the importance of communicative function and speaker intentions over fixed noun categories (Wisniewski et al., 2003). Finally, this accords with findings from psycholingistics research (Barner & Snedeker, 2005, 2006; Fieder et al., 2014), and from comprehension and production studies that show that English speakers rely on syntactic cues like quantifiers and determiners for interpreting countability in context (Chiarelli et al., 2011; Fieder et al., 2014, 2018).

Many nouns thus do not fall into neat, distinct categories of count and non-count. They form a continuum from strongly countable to strongly non-count, with many flexible nouns and multiple gradations of flexibility in between. There seems to be a larger number of strongly countable nouns on the right end of the spectrum than elsewhere, which reflects the more default nature of nouns, which is to refer to known objects in our environment. The so-called articles *a/an* and *the* or their absence function to delimit the potential range of meaning of a noun phrase to, say, a specific, individuated object, a mass, an abstract term, or a specific instance thereof (K. Lee, 2017, 2021a). Thus, the term *delimiter* may be a more accurate term to describe their function, while the term *article* can be ambiguous or confusing (especially, say, in a writing course, when one also refers to published texts as articles).

While it may be convenient to present simplistic categorizations for beginning levels, more intermediate and advanced levels need to learn countability as a gradient and context-dependent concept that depends on a speaker's intended meaning. Students need to learn about strongly count, flexible, and strongly mass usage, and the semantic shifts and differences in meaning or nuance that arise from using flexible nouns in different ways. For example, one drinks *coffee* or *water*, and can order *a coffee* or sometimes *a water*; one can buy *chicken* as a prepackaged food material or buy *a chicken*, i.e., a whole chicken. Then other semantic alternations need to be taught. For example, one can speak of *a penguin*, or of *penguins* more generically, i.e., all penguins or penguins in general. Abstract nouns are usually used as bare singular non-count nouns, but can be made more countable if discussing a specific instance, type, or example of that idea or property (e.g., *democracy* vs. *the democracy of the ancient Greeks*).

Learners can be taught some heuristic principles for thinking about how to use a noun in a sentence. This can be helpful when they make errors, when providing them feedback, or when they encounter difficulties articulating well-formed noun phrases. A few example heuristics are shown in Table 4, which derive from this study and from other related research (namely, Lee, 2017, 2021a). This can be helpful for teaching students how to use a noun, or for guiding them to understand errors that they have made.

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

Table 4
Heuristic Questions for Learners

Heuristic	Example
Is it a thing or an object? Is it a material?	chair cf. wood glass cf. a glass
Is it a material, or one portion, or one type, of such material?	salad cf. a salad coffee cf. a coffee
Is it a more abstract term? Or does it refer to a specific instance, example, or type of what is normally more abstract?	freedom cf. the freedom to choose jazz cf. the jazz of New Orleans
Does it refer to a general activity, or a specific instance of an activity, i.e., an event or occurrence?	I like hiking cf. a great hike
Does it refer to a general description, condition or property, or one specific instance, example, or type thereof?	redness cf. the redness of her skin pain cf. the pain of not knowing
Does it refer to one specific thing, or a set or group of multiple items?	an egg cf. eggs a vehicle cf. vehicles
Does it refer to one specific thing, or that sort of thing in general?	a cheetah cf. cheetahs a chair cf. chairs

The primary pedagogical value of this study lies in its potential to aid in the selection and use of suitable nouns and grammatical patterns for EFL lessons and materials. The BNCI and GNCI provide means for identifying nouns that differ in their actual countability and usage, so that educators can more readily find and select suitable nouns for lesson materials. Educators or materials developers can find nouns by scale ratings, or by deciles (the scale ratings are broken into deciles for sorting and finding more nouns in a given range of countability). They can refer to noun indices in the GNCI or BNCI to see the likelihood of a noun being strongly countable (near 1.0 on the scale), strongly non-countable (near 0 on the scale), or flexible (around the middle of the scale). The intermediate levels (by decile or by GNCI or BNCI scores) indicate nouns that are more polysemous or semantically variable, e.g., *glass* (GNCI index: 0.467, 4th decile) can refer to a material, or to a drinking vessel. Such nouns deserve more careful and systematic attention, especially at more intermediate and advanced levels. Users of these indices can sort nouns by index values or deciles, and by lexical frequency, that is, by raw frequency of token counts in the COCA dataset, by frequency deciles based on COCA, or by the frequencies from Brysbaert et al. (2014), which are also in the dataset. The data are available online.[2]

Nouns that are higher frequency and more typically mass or count may be used to introduce the common patterns, e.g., more typically countable nouns like *people, workers, patient, book, door, story*, and *process*, and more typically mass nouns like *peace, information, education, honey, food*, and *music*. For more intermediate lessons, more common flexible nouns with intermediate index values or decile ranks can be used as examples of contextual variation in countability. The particular contexts and differing meanings or nuances can be discussed for nouns like *state, government, kind, effect, force, time, place, city,* and *night*. Teachers can find actual contextual examples (if not from a corpus or from their own examples, then from an online search for the target nouns to find online examples used in context), and students can be asked to infer the nuances and why the nouns are so used in those contexts. For more advanced lessons, less frequent vocabulary from the mass, count, and flexible categories can be chosen. Also, less frequent examples of nouns can be chosen for lesson materials, such as those in the low-intermediate and high-intermediate deciles or index values. For example, *water* (Decile 2) is most often mass, but can be used when ordering a portion (ordering a water, i.e., a glass or bottle of water) or when referring to a type or batch of water (e.g., *the water on Mars is toxic*, or *the water in the pond is too deep*); *name* (Decile 7) is usually countable (a person's name), but can be more abstract (e.g., *a leader in name only*, implying "not a real leader"). Examples of such expressions can be found online with expanded Google search syntax (e.g., "+*in name only*" in quotation marks searches for that specific string, and the plus sign forces a search with minor words). After finding nouns from a countability index, teachers can search online for examples with adjectives, determiners, quantifiers, plural markings, prepositional phrases, and other patterns. Of course, finding examples from authentic materials will be preferable instead of, or as a supplement to, standard teaching materials.

[2] Sample sets of nouns with GNCI, BNCI, and other data will be made available at https://www.enwiki.org/w/count_nouns and https://www.enwiki.org/w/flexible_nouns. The full data sets of 20,000 nouns will be available at the author's GitHub repository at https://www.github.org/kentlee7/nouns.

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

In such manner, English noun phrase patterns can be presented in a more systematic and intuitive manner. Learners and educators should view traditional rules in their textbooks circumspectly, and likewise, the simplistic count / non-count classifications in some of their dictionaries. However, if students learn and focus on countability as a gradient, or as a choice that depends on the context and intended meaning, the choice of definite, indefinite or bare noun phrases may make more sense and seem less intimidating.

CONCLUSION AND IMPLICATIONS

Clearly much more detailed research is needed. The data set yields reliable data for language teachers and materials developers who need relatively straightforward information on nouns that are commonly used as count, non-count, or flexible nouns. However, the data are not sufficiently reliable for more detailed linguistic and statistical analysis, or addressing more specific research questions. Errors in COCA have been less problematic for other published studies, but for this topic, very high accuracy is needed in tagging. COCA and similar corpora are also not parsed, e.g., for count and non-count nouns. It is thus difficult to extract noun data that are accurate enough for more detailed analysis. Better corpus data would allow for studying issues such as first and subsequent mention of nouns, noun usage that does not follow the first/subsequent mention rule, superordinate nouns (e.g., *furniture*), other nouns that do not follow traditional rules, other more unusual delimiter (article) usage patterns, and more about the degree of spread of the countability spectrum and types of nouns along different areas of the spectrum. More detailed analysis of noun usage patterns across written and spoken genres, and of compound noun patterns, is also needed.

Since COCA consists of data from 1990 to 2019, it may be outdated for a few noun usage patterns that have changed since then. For example, *text* serves readily as a mass or count noun, but in recent years the countable use (a text message) has become more commonplace. COCA and similar corpora use the CLAWS tagging system, with an accuracy of approximately 96–97% (Leech & Smith, 2000), leading to a number of incorrectly tagged tokens in a large corpus. CLAWS and other taggers can be error prone in general (Rayson et al., 2007), but especially so for spoken corpora due to disfluencies, overlaps, pragmatic markers, audio quality, and other issues (Galiano & Semeraro, 2023). Since COCA is also not fully parsed, noun phrases or syntactic relations must often be inferred indirectly. The fact that some tokens are deliberately obscured for watermarking purposes can also introduce errors or difficulties for researchers. For more accurate corpus analysis of noun phrase patterns, a more reliable corpus with more reliable tagging and parsing would be preferable. While the results of this study can be useful for language teachers, more recent and more reliably tagged and parsed data sets would provide better data, not only for language education, but also for applied linguistics research on noun phrase usage.

Finally, this does not attempt to address context-specific issues such as associative nouns, e.g., when mention of *the office* allows for mention of *the secretary* or *the photocopier*. It does not address other quirks, such as plural-only noun meanings (e.g., *goggles*), other well-known ambiguities (e.g., *at school, in the hospital*), proper nouns, acronyms, or compound noun phrases.

English nouns do not fall into neat, discrete categories of count nouns and non-count nouns. They form a continuum, with strongly mass or non-count nouns on one end, a large number of strongly count nouns on the other end, and a large number of flexible nouns spread fairly evenly in between. This general finding leads to many more research questions that require much further research with better data.

Teachers and materials developers can present noun usage as a gradient or continuum from strongly non-count to strongly countable, with many flexible nouns in between that can vary in their nuance, depending on how they are used. Educators can use these countability indices to find nouns that are more mass-like, countable, or flexible, and to find nouns that would be suitable for different types of lessons. Students will be better served by learning the right questions to ask of each noun in context, instead of relying on traditional rules that are difficult to apply in real contexts.

References

Amuzie, G. L., & Spinner, P. (2013). Korean EFL learners' indefinite article use with four types of abstract nouns. *Applied Linguistics*, *34*(4), 415–434. https://doi.org/10.1093/applin/ams065

Antes, T. A. (2019). Countability versus flexibility: Rethinking vocabulary and determiner instruction. *TESL-EJ*, *23*(2), 1–18.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (release 2)* [Dataset]. Linguistic Data Consortium.

Baldwin, T., & Bond, F. (2003a). A plethora of methods for learning English countability. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing* (pp. 73–80). https://aclanthology.org/W03-1010/

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

Baldwin, T., & Bond, F. (2003b). *Learning the countability of English nouns from corpus data*. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics* (pp. 463–470). Association for Computational Linguistics. https://doi.org/10.3115/1075096.1075155

Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, *97*(1), 41–66. https://doi.org/10.1016/j.cognition.2004.06.009

Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: Individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development*, *2*(3), 163–194. https://doi.org/10.1207/s15473341lld0203_2

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Chiarelli, V., El Yagoubi, R., Mondini, S., Bisiacchi, P., & Semenza, C. (2011). The syntactic and semantic processing of mass and count nouns: An ERP study. *PloS One*, *6*(10), e25885. https://doi.org/10.1371/journal.pone.0025885

Cho, Sookhee. (2006). Korean EFL learners' judgments of countability of English nouns in isolation and context. *Korea Journal of English Language and Linguistics*, *6*(4), 869–893.

Choi, S. H., & Ionin, T. (2021). What's the smallest part of spinach? A new experimental approach to the count/mass distinction. *Experiments in Linguistic Meaning*, *1*, 113–124. https://doi.org/10.3765/elm.1.4867

Choi, S. H., Ionin, T., & Zhu, Y. (2018). L1 Korean and L1 Mandarin L2 English learners' acquisition of the count/mass distinction in English. *Second Language Research*, *34*(2), 147–177. https://doi.org/10.1177/0267658317717581

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, *14*(2), 159–190. https://doi.org/10.1075/ijcl.14.2.02dav

Davies, M. (2013). Google Scholar and COCA-Academic: Two very different approaches to examining academic English. *Journal of English for Academic Purposes*, *12*(3), 155–165. https://doi.org/10.1016/j.jeap.2013.01.003

Drożdż, G. (2020). New insights into English count and mass nouns—the cognitive grammar perspective. *English Language & Linguistics*, *24*(4), 833–854. https://doi.org/10.1017/S1360674319000480

Fieder, N., Nickels, L., & Biedermann, B. (2014). Representation and processing of mass and count nouns: A review. *Frontiers in Psychology*, *5*, Article 589. https://doi.org/10.3389/fpsyg.2014.00589

Fieder, N., Nickels, L., Krajenbrink, T., & Biedermann, B. (2018). Garlic and ginger are not like apples and oranges: Effects of mass/count information on the production of noun phrases in English. *Quarterly Journal of Experimental Psychology*, *71*(3), 717–748. https://doi.org/10.1080/17470218.2016.1276203

Galiano, L., & Semeraro, A. (2023). Part-of-speech and pragmatic tagging of a corpus of film dialogue: A pilot study. *Corpus Pragmatics*, *7*(1), 17–39. https://doi.org/10.1007/s41701-022-00132-9

Gillon, B. S. (1996). The lexical semantics of English count and mass nouns. In E. Viegas (Ed.), *Breadth and depth of semantic lexicons* (pp. 51–61). Kluwer Academic.

Grimm, S., & Wahlang, A. (2021). Determining countability classes. In T. Kiss, F. J. Pelletier, & H. Husić (Eds.), Things and stuff: The semantics of the count-mass distinction (pp. 357–376). Cambridge University Press. https://books.google.co.kr/books?id=bDMyEAAAQBAJ

Inagaki, S. (2014). Syntax–semantics mappings as a source of difficulty in Japanese speakers' acquisition of the mass–count distinction in English. *Bilingualism: Language and Cognition*, *17*(3), 464–477. https://doi.org/10.1017/S1366728913000540

Jespersen, O. (1924). *The philosophy of grammar.* University of Chicago Press.

Katz, G., & Zamparelli, R. (2012). Quantifying count/mass elasticity. In J. Choi, E. A. Hogue, J. Punske, D. Tat, J. Schertz, & A. Trueman (Eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics* (pp. 371–379). Cascadilla Proceedings Project. https://lingref.com/cpp/wccfl/29/paper2723.pdf

Kim, Dohee, & Kim, Hyunwoo. (2026). Effects of L1 knowledge and task type on bilingual acquisition and processing of English count and mass nouns. *International Journal of Bilingualism*, *30*(3), *939–953*. https://doi.org/10.1177/13670069251360171

Kim, S. (2018). A lexicographic approach to teaching the English article system: Help or hindrance? *Lexikos*, *28*, 196–220. https://doi.org/10.5788/28-1-1462

Kiss, T., Pelletier, F. J., Husić, H., Simunic, R. N., & Poppek, J. M. (2016). A sense-based lexicon of count and mass expressions: The Bochum English Countability Lexicon. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2810–2814). European Language Resources Association. https://aclanthology.org/L16-1447/

Krifka, M. (1995). Common nouns: A contrastive analysis of Chinese and English. In G. N. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 398–411). University of Chicago Press.

Lee, Hikyoung. (1999). Variable article use in Korean learners of English. *University of Pennsylvania Working Papers in Linguistics*, *6*(2), 35–47.

Lee, K. (2017). A "the" or the "a"? L2 learner problems and patterns. *Korea TESOL Journal*, *13*(2), 25–48.

Lee, K. (2021a). A study on the viability of a cognitive approach to teaching articles. *Studies in English Education*, *26*(1), 29–55. http://doi.org/10.22275/SEE.26.1.02

Lee, K. (2021b). Teaching English articles: Addressing university students' misconceptions. *Journal of Learner-centered Curriculum and Instruction*, *21*(2), 1–25. http://doi.org/10.22251/jlcci.2021.21.2.1

Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: The tagging of the British National Corpus. *COLING 1994 Volume 1: The 15th International Conference on Computational Linguistics*. Lancaster University.

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)

Leech, G., & Smith, N. (2000). *Manual to accompany the British National Corpus (Version 2) with improved word-class tagging*. University Centre for Computer Corpus Research on Language (UCREL), Lancaster University. https://ucrel.lancs.ac.uk/bnc2/

Ma, Yumi, & Oh, Yoonja. (2014). Korean EFL learners' perceptions and uses of count/mass nouns. *English21*, *27*(4), 463–481. http://doi.org/10.35771/engdoi.2014.27.4.021

MacDonald, D., & Carroll, S. E. (2018). Second-language processing of English mass-count nouns by native-speakers of Korean. *Glossa*, *3*(1), Article 46. https://doi.org/10.5334/gjgl.363

McCawley, J. D. (1975). Lexicography and the count-mass distinction. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, *1*, 314–321. https://doi.org/10.3765/bls.v1i0.2335

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, *3*(4), 235–244. https://doi.org/10.1093/ijl/3.4.235

Min, Soo-Jung. (2014). Countability of mass nouns among the EFL learners in Asian context. *Journal of Studies in Language*, *30*(1), 91–106. http:/doi.org/10.18627/jslg.30.1.201405.91

Payne, J., & Huddleston, R. (2002). Nouns and noun phrases. In R. Huddleston & G. K. Pullum (Eds.), *The Cambridge grammar of the English language* (pp. 323–524). Cambridge University Press. https://doi.org/10.1017/9781316423530.006

Pelletier, F. J. (1975). Non-singular reference: Some preliminaries. *Philosophia*, *5*(4), 451–465. https://doi.org/10.1007/BF02379268

Pelletier, F. J. (2012). Lexical nouns are both+ MASS and+ COUNT, but they are neither+ MASS nor+ COUNT. In D. Massam (Ed.), *Count and mass nouns across languages* (pp. 9–26). Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199654277.003.0002

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.

Rayson, P., Archer, D. E., Baron, A., Culpeper, J., & Smith, N. (2007). Tagging the Bard: Evaluating the accuracy of a modern POS tagger on Early Modern English corpora. *Proceedings of the Corpus Linguistics Conference: CL2007*. Lancaster University.

Rijkhoff, J. (2008). On flexible and rigid nouns. *Studies in Language*, *32*(3), 727–752. https://doi.org/10.1075/sl.32.3.12rij

Smith-Stark, T. C. (1974). The plurality split. *Chicago Linguistic Society*, *10*(1), 657–672.

Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 46–58. https://doi.org/10.1037/0278-7393.28.1.46

Wisniewski, E., Lamb, C., & Middleton, E. (2003). On the conceptual basis for the count and mass noun distinction. *Language and Cognitive Processes*, *18*(5–6), 583–624. https://doi.org/10.1080/01690960344000044

Xue, M. (2010). Countable or uncountable? That is the question—Lexicographic solutions to nominal countability in learner's dictionaries for production purposes. *Lexikos*, *20*(1), 540–558. https://doi.org/10.5788/20-0-155

Zamparelli, R. (2020). Countability shifts and abstract nouns. In F. Moltmann (Ed.), *Mass and count in linguistics, philosophy, and cognitive science* (pp. 191–224). John Benjamins. https://doi.org/10.1075/lfab.16.09zam

www.kci.go.kr

© 2026 The Modern English Education Society (MEESO)