A Comparative Analysis of Evidentiality in L1 And L2 Academic Writings and Its Reflected Identity Construction

Li Xin[1*], Li Yang[2]

[1]*School of Foreign Languages and Literature, Shandong University, Jinan, Shandong，China;*
[2]*Health Management Center, the Second Hospital of Shandong University, Jinan, Shandong, China.*
Corresponding Author.

Abstract: Evidentiality is a linguistic phenomenon in which speakers expose the source of their knowledge and convey their attitudes towards knowledge. Data for our study are composed of written texts including L2 masters' thesis and L1 published research articles and are analyzed through the software *AntConc* to explore characteristics of the use of evidentiality and identity construction of L1 and L2 writers. It is found that there is an overall similarity but still considerable disparities between L1 and L2 writers in their specific uses of evidentials. Moreover, due to differences in ways of thinking and cultural consideration, identity constructions by L1 and L2 writers are also different.

Keywords: Evidentiality; Academic Writing; Identity Construction

1. Introduction

Evidentiality is a linguistic phenomenon in which speakers expose the source of their knowledge and convey their attitudes towards knowledge. It is usually expressed by "evidentials" or "evidential markers" [1, 2, 3]. Two types included in evidentiality are morphological evidentials and lexical evidentials. the former refers to the prefixes and suffixes of verbs while the latter refers to adverbs, adjectives, verbs, modality verbs, etc. This study investigates the distinction between L2 masters' theses and L1 experts published articles in terms of lexical evidentials. It is feasible for writers to reasonably allocate the structure of their writings and efficiently convey their viewpoints through evidentiality.

However, the comparison of the use of evidentiality between L2 students' thesis writing and L1 native published research articles have not undergone in-depth analysis. Furthermore, it is critical to draw on corpus to increase master students' proficiency in academic writing in English as writing in foreign language is a core requirement for English-major students and it is hard to develop in a short period of time if correct instructions are ignored.

2. Literature Review

2.1 Evidentiality

It is generally believed that evidentiality can be interpreted from its narrow sense and broad sense. In a narrow sense, evidentiality could be used to investigate the source of knowledge, highlighting the grammatical aspect of the source of a proposition. In this way, evidentials are regarded as a group of suffixes used to express the source and degree of certainty of knowledge [4]. In a broad sense, evidentiality addresses directly the source and attitude of speaker's knowledge [2].

Previous studies on evidentiality involve various types of discourse, including academic discourse [5], news discourse [6], advertising discourse, and court discourse [7]. Recent studies of evidentiality on web anchors, live broadcast and pandemic discourse have also evolved. Although previous studies have looked into evidentiality in academic discourse, they have not further refined the use of evidentiality in different sections of academic writings, particularly to analyse disparities between L1 experts' published articles and L2 masters' theses.

Proper knowledge of the use of evidentiality is especially significant for masters to improve the accuracy and logicality of their writings. the investigation of the different use of evidentiality between non-native masters and native experts sheds light on the discussion and conclusion section written by non-native masters.

2.2 Identity Construction

Identity entails the meaning and experience of subjectivity, aiding us in understanding who we are and how to be. Instead of being static and fixed, it is temporary and procedural in nature. Self-identification, identity work, and identity regulation are three major factors that affect identity construction. Self-identity focuses on how a person sees themselves. Identity work refers to that active process that creates the self-identity. Identity regulation means the social and organisational processes that could influence how a person acts [9].

It is inevitable for authors to elaborate on the information relevant to their researches such as their findings and significance, indicating that their identities are constructed on the basis of completing these tasks [10]. In addition, identity is a specific trait formed by people in the process of interaction. Identity is constantly built, negotiated, and reproduced in this dynamic process when the author engage readers into the presentation of academic discourse [9]. the author's attitude and behaviour are presented in this ever-changing process where the identity is also built [11]. Thus, identity construction of author's identity in academic writings has been one of the focus of identity research, mainly exploring discourse strategies for authors to construct their identities.

In academic writings, it is important for authors to forge specific relationships and interactions with readers and construct different identities in the interaction to better coordinate their writing resources and purposes - introducers and evaluators. This study is guided by the following questions: What are the characteristics of the evidentiality used by L1 Chinese masters and L1 English experts; What similarities and disparities are discernible in the process; If differences exist, what are the factors that contribute to them?

3. Method

3.1 Corpora

Data for our study are composed of written texts including L2 masters' thesis and L1 published research articles. One of the sub-corpus, masters' thesis writing corpus (MTWC), is established by selecting masters' thesis in the discipline of linguistics from CNKI and WANFANG DATA from 2014 to 2022.135 thesis are randomly selected with a total number of 579294 words. It should be noted that the masters' thesis collected on the two data platforms are excellent theses, representing a relatively high level of writing among masters' students. Detailed information on article choice is provided in the following **Table 1**.

Table 1. Detailed Information about Two Sub-Corpora

Corpus	L1 Chinese	L1 English
Sources/Journals	CNKI WANFANG DATA	System Applied Linguistics English for Specific Purpose
Year	2014-2022	2014-2022
Text type	MA's thesis	Published articles
No. of texts	135	298
Total number of words	579294	588800

3.2 Data Analysis

There is no single, unified criteria for the classification of evidentials. Based on the characteristics of academic writings and the classification of evidentials proposed by previous studies [2, 3, 12], this paper divides evidentials into the following five types and tends to reveal the similarities and differences used by L2 masters and L1 experts. This study annotates all the evidentials in the two sub corpora (MTWC and EPWC) by combining *AntConc* with manual screening, particularly excluding those with polysemy. Then, the standard frequency of the use of evidentials is calculated in the two sub corpora, and the Chi square Calculator is used to run a significance test on them. As **Table 2.**

Table 2. the Taxonomy and Function of Evidentials

Categories	Functions	Forms		Examples
Reporting Evidential	citing viewpoints to ensure fairness and objectivity	reporting verbs		X said/reported/argued/emphasized/revealed...
		non-verbal forms	prepositional phrases	according to X, in X's opinion/view...
			noun+that	claim/fact/observation/finding/view that...
			It-clause	It is suggested/argued that...

		Parenthesis citation		(X, 2002)
Sensory Evidential	an intuitive way to perceive the world	Sensory verbs (visual, auditory, or other senses)		I/we (can) see that..., as we can see...,I/we (can) hear that, It smells/tastes that..., we touch/feel...
Hypothetical Evidential	prefabricating the grow path of the matter	If-clause		If...
Conviction Evidential	conducting subjective speculation	psychological verbs		I/we think/suggest/believe
		non-verbal forms	prepositional phrases	in my view/opinion, to my knowledge
			psychological nouns+that	my view/opinion is that...
			It-clause	It can be suggested, it may be argued
			adverbs	arguably, admittedly
Inferential Evidential	evaluating the verifiability of claims	verbs	epistemic modality verbs	must, should, can, could, may, might, will, would
			seem-like copula	seem/appear to, it seems that...
		non-verbal forms	nouns	possibility, certainty
			It-clause	It is obvious/evident/possible/clear that...
			modal adverb	perhaps, possibly, maybe, seemingly, certainly

4. Results and Discussion

4.1 General Comparison of Evidentials
Different distributions of the use of evidentials in MTWC and EPWC are presented in **Table 3**. *Figure 1* is given based on the data in table 3 to more intuitively observe the distinction of their distributions in terms of their standard frequency. **Figure 1** shows that the types employed in NP-, PP-, or VP-based bundles in MTWC are substantially greater than those used in EPWC.

Table 3. Different Distributions of Evidentials in MTWC and EPWC

*=significant p < 0.05; **=significant p < 0.01

Categories of Evidentials			MTWC(579294)			EPWC(588800)		
			Actual frequency	Standard frequency	proportion	Actual frequency	Standard frequency	proportion
Reporting Evidential	reporting verbs		146	0.252	1.369%	141	0.239	0.940%
	non-verbal forms	prepositional phrases*	578	0.998	5.421%	174	0.296	1.161%
		noun+that	177	0.306	1.660%	263	0.447	1.754%
		It-clause	27	0.047	0.253%	8	0.014	0.053%
	Parenthesis citation		1699	2.933	15.935%	5350	9.086	35.683%
	Total*		2627	4.535	24.639%	5936	10.082	39.593%
Sensory Evidential	Sensory verbs*		321	0.554	3.011%	205	0.348	1.367%
Hypothetical Evidential	if-clause		387	0.668	3.630%	450	0.764	3.001%
Conviction Evidential	psychological verbs		16	0.028	0.150%	141	0.239	0.940%
	non-verbal forms	prepositional phrases	3	0.005	0.028%	5	0.008	0.033%
		psychological nouns+that	0	0.00	0.000%	0	0.000	0.000%
		It-clause*	1	0.002	0.009%	19	0.032	1.267%
		adverbs	6	0.010	0.056%	32	0.054	0.213%

		Total*	26	0.045	0.244%	197	3.350	1.314%
Inferential Evidential	verbs	epistemic modality verbs*	6631	11.447	62.193%	7170	12.177	47.822%
		seem-like copula	238	0.411	2.232%	554	0.941	3.695%
	non-verbal forms	nouns	131	0.226	1.229%	113	0.192	0.754%
		It-clause	150	0.259	1.407%	89	0.151	0.594%
		modal adverb	151	0.261	1.416%	279	0.474	1.861%
		Total*	7301	12.603	68.477%	8205	13.935	54.726%
Total			10662	18.405	100%	14993	25.464	100%
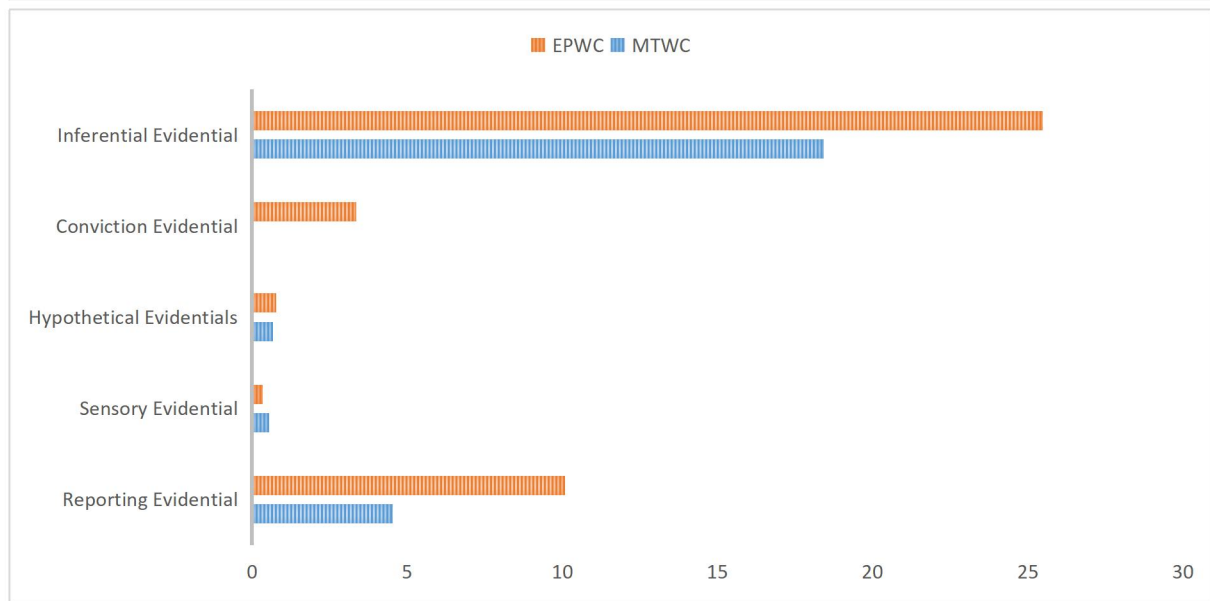
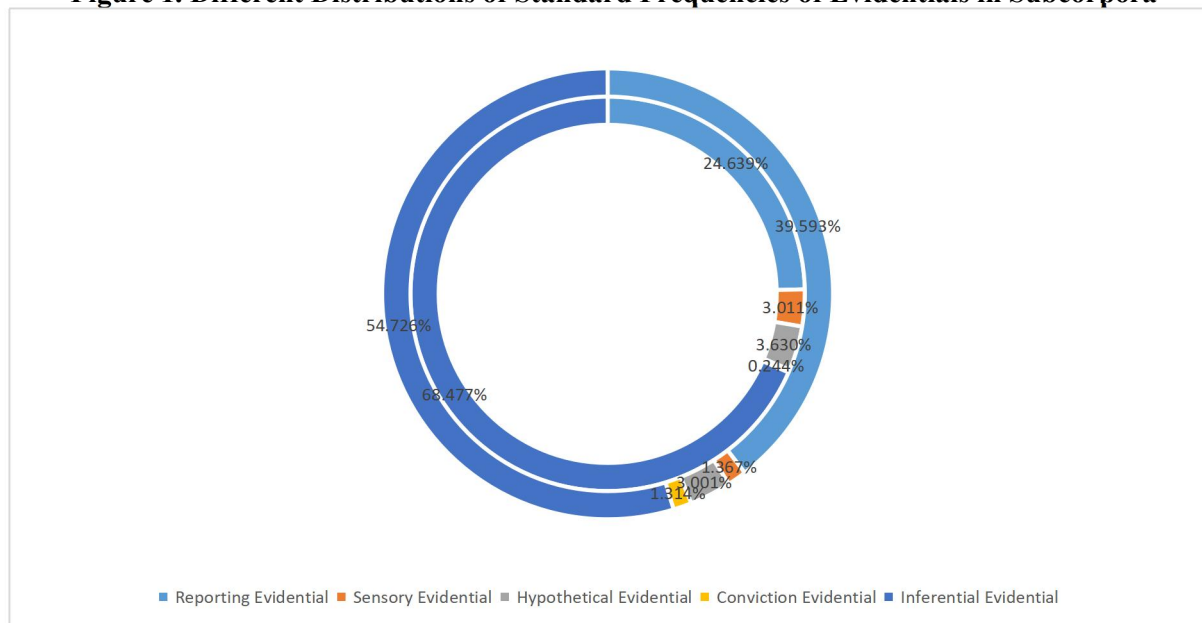**Figure 1. Different Distributions of Standard Frequencies of Evidentials in Subcorpora**

Figure 2. Different Proportions of Evidentials in MTWC and EPWC

Figure 2 shows different proportions of evidentials in MTWC and EPWC. It can be seen that proportions of evidentials in MTWC and EPWC are roughly similar. the highest proportion of the category on both MTWC and EPWC is inferential evidential, followed by reporting evidential. the other three evidentials, namely, conviction evidential, sensory

evidential, and hypothetical evidential, have a relatively small proportion in both sub-corpora, with the lowest proportion being conviction evidential. It demonstrates that both L1 and L2 writers are aware of the use of evidential in a similar way. Both realize that evidentiality reflects the communication rules of the academic community - the agreed academic research norms and language display forms. the rational use of evidentiality not only reflects the respect and flexibility with which academic research norms are applied, but it is also an important means of building academic identity.

However, it is found that despite the similar distribution of evidentials used by MTWC and EPWC, their specific proportions and modes of application vary widely. For example, the proportion of inferential evidential in MTWC is much higher than that in EPWC, while its share of reporting evidential is lower than that in EPWC. Moreover, although conviction evidential accounts for the smallest proportion in both sub-corpora, its share in EPWC is higher than that in MTWC. It reflects that despite little variation in the overall use of evidentials between L1 and L2 writers, there is a distinction in their frequency and specific proportions, resulting in variations in the way the L1 and L2 writers construct their own identities.

4.2 Detailed Analysis of Different Distributions of Evidentials and its Reflected Identity Construction

L1 writers employs reporting evidential more frequently than L2 writers, and its proportion to total evidentials is also greater than that of L2. Reporting evidential refers to the direct use of existing knowledge and its paraphrase. Information could be obtained through experience either directly or indirectly. the use of reporting evidential enhance the credibility and objectivity of their academic writings, making them accessible to readers.

Although parenthesis citation is used with the highest frequency and proportion in reporting evidential in both sub-corpora, L1 has a substantially higher frequency and proportion than L2. Writers avoid conveying personal attitudes and tendencies by using parenthesis citation to express the viewpoints of other professionals. the academic writing has been proved to be objective by citing other studies,

minimizing the subjective undertone to prevent its own responsibility and unnecessary disputes. As seen in (1):

(1)... other comments (17-19) provide evidence of the fluidity in the structure of the RA, a focus on the aesthetic, a balance between explanatory text and mathematical reasoning (5-7) (McGrath & Kuteeva, 2012)... (EPWC)

L2 writers have a higher frequency and proportion of sensory evidential compared to L1 writers. People's cognition and perception of the world, including vision, hearing, smell, and touch, are directly derived from their sense-based experience. It is the result of human induction of natural and social phenomena [6]. Sensory evidential is primarily represented through visual and auditory means. Sensory evidence is used to address the fact that information originates from the firsthand experience by providing relevant evidence. In EPWC, L1 writers are accountable for the credibility of the information source. Through the use of sensory verbs, notably "see", L1 writers tend to capture readers' attention, lead them to concur with their own thoughts and opinions, and establish their identity as a conductor.

EPWC has a substantially larger frequency and proportion of hypothetical evidential than MTWC. Hypothetical evidential, often realized in the form of "if-clause", differs from reporting evidential and conviction evidential in that the speaker has neither personally experienced nor collected information from people or other sources. Instead, it is a reasonable fiction created by the writer based on their own experience and subjective knowledge, and the inferences reached from the assumption may not necessarily conform to reality [6].

Despite the fact that both corpora have a very modest fraction of conviction evidential, it is evident that the frequency and proportion of conviction evidential in EPWC are higher than those in MTWC. the most commonly used form, in both sub-corpora, is psychological verb, followed by adverb and it-clause, but they rarely use the form of prepositional phrases and psychological noun. the relatively lower use of conviction evidential reflects the features of academic writings. Conviction refers to the viewpoint, imagination and speculation of writers. Writers use conviction evidential to express their subjective opinions

or ideas. In this process, L1 writers invariably produce discourse that is tied to their personal views or opinions, lending subjectivity to the entire discourse production. As seen in (2):

(2) *I think* that this kind of tool requires a change of perspective in the students and we, as their teachers, should facilitate it. (EPWC)

The frequency and overall proportion of conviction evidential used by L2 writers are relatively low. Even if there is conviction evidential in their writings, the expression of uncertainty such as "should, could" almost follow it, indicating that L2 writers rarely involve subjective opinions in their academic writings. Despite the occasion existence of subjective expressions, the concept of possibility will be used to to lessen them in order to prevent subjectivity from affecting the legitimacy and objectivity of the overall investigation. Thus, L2 writers' role as evaluators are hidden. In contrast, L1 writers use conviction evidential more frequently, the majority of which are produced by using psychological verbs such as "I think, I believe". In this process, L1 writers construct their identity as evaluators. For example, in (3):

(3) On the contrary, I think that the Westerners go shopping only in supermarkets. (EPWC)

Inferential evidetial accounts for the largest part in both corpora. It is used more frequently in EPWC than that in MTWC, but the share of inferential evidential is higher. Apart from the subcategory of "it-clause", the other subcategories of inferential evidential are used more frequently in EPWC, but the most frequently used subcategory in both corpora is epistemic modality verb. Through inferential evidential, writers convey their opinions regarding the possibility and status of facts and highlight their attitudes and stance towards the results of current work or previous studies, further cementing their position as an evaluator.

4.3 Factor Analysis on Evidential Differences

Firstly, the contrasts in thoughts between China and the West result in differences inherent in the writing itself. the Chinese, with a "holistic view", place a high value on overall appraisal and systematic grasp of matters, while the West emphasize "analytical" thinking, which is reflected in disparities in writings. Chinese emphasizes parataxis while English emphasizes hypotaxis. English

depends on formal techniques to indicate the logical or semantic linkages between words or sentences, which bears the features of explicit discourse markers. In contrast, Chinese contains the features of implicit discourse markers, which leverage internal semantic connection to convey implicit meanings. Therefore, the awareness of using evidential among Chinese people is not as strong as in the West, resulting in a higher overall frequency of evidentials used by L1 writers compared to L2 writers.

Moreover, due to cultural considerations, Confucianism-influenced Chinese pay attention to euphemism and modesty of language and attempt to avoid using expressions that contradict one another while convey their academic views. With a compromise approach adopted, Chinese typically emphasize objectivity and minimize the intervention of their own subjective thoughts. Despite the appreciation of objectivity and credibility in academic writings in the West, direct presentation of opinions is also valued, with a strong subjective undertone. Furthermore, there is a difference in the level of semantic knowledge between China and the West. It can be split into high context culture and low context culture based on the varied degrees of dependence of communication on the environment in different cultures [13]. China is a high context culture, in which readers are sensitive to context and can capture the potential meanings inherent in their surroundings. the West, on the other hand, is a low context culture that demands interpreting the meaning through already conveyed words and relying on the language itself for interactive communication. As a result, L2 writers tend to use reporting and inferential evidential to help readers understand their potential intentions, while L1 writers, in addition to using the above two evidentials, also draw on conviction evidential to directly express their views, to aid readers in perceiving their situations, attitudes, or stance. Therefore, L2 writers are more likely to play a role as an introducer and an evaluator while L1 writers prefer to construct their identity as an evaluator and a conductor.

5. Conclusion

This study investigates the characteristics of employing Evidentiality in academic writing

by L1 experts and L2 masters based on the five classifications of Evidentiality as well as the characteristics of their identity formation.

Based on the taxonomy of evidentiality, this study investigates the characteristics of employing evidentiality in academic writings by L1 experts and L2 masters and the features of their identity construction. It is found that there is an overall similarity in the use of evidentials by L1 and L2 writers. For example, in the two sub-corpora, inferential evidential is used most frequently and accounts for the largest proportion, followed by reporting evidential, while hypothetial and sensory evidential are quite seldom used by both. However, there are still considerable disparities between L1 and L2 writers in their specific uses of evidentials. In addition to using evidentials more frequently than non-native speakers in general, native speakers also use a higher percentage of conviction evidential. Due to differences in ways of thinking and cultural consideration, the features in the application of evidentials also demonstrates the differences in identity construction by L1 and L2 writers. Non-native masters are more likely to use inferential evidential and reporting evidential to establish their identity as introducers and evaluators while native experts prefer to use inferential evidential, reporting evidential and conviction evidential to construct their identity as evaluators and conductors.

However, there are also some flaws in this study. This study exclusively compares academic writings in the field of applied linguistics, and the size of corpus has some restrictions. Future studies will be able to compare academic writings from different fields within a broader corpus.

References

[1] Sapir, E. (1921). Language. An introduction to the study of speech. New York: Harcourt, Brace &Co.

[2] Chafe, W. (1986). Evidentiality in English conversation and academic writing. In: W. Chafe and J. Nichols, (Eds.), Evidentiality: the linguistic coding of epistemology. Norwood, N J: Ablex. pp. 261-272.

[3] Hu, Z. (1994). Provability, news reporting and argumentative language. Foreign Language Studies, 2, 22-28.

[4] Boas, F. (1947). Kwakiutl grammar, with a glossary of the suffixes. Transactions of the American Philosophical Society, 37:201-377.

[5] Yang LX. (2015). Identity construction in English research articles: An evidential perspective. Foreign Language Education, 36(2), 21-25.

[6] Wang GF & Yu XY. (2011). A typological study of the evidential in Chinese-English news text. Journal of Xi'an International Studies University, 19(2):30-33.

[7] Luo GH & Liao MZ. (2013) Evidentiality in Courtroom Discourse. Studies in Language and Linguistics, 33(4):92-95.

[8] Alvesson, M., Lee Ashcraft, K., & Thomas, R. (2008). Identity matters: Reflections on the construction of identity scholarship in organization studies. Organization, 15(1), 5-28.

[9] Sveningsson, S., & Larsson, M. (2006). Fantasies of Leadership: Identity Work. Leadership, 2, 203-224.

[10] Dickerson, P. (1996). Let me tell us who I am: the discursive construction of viewer identity. European Journal of Communication, 11(1), 57-82.

[11] Wooffitt, R. (1992). Telling tales of the unexpected: the organization of factual discourse. Rowman & Littlefield.

[12] Brugman, C. M., & Macaulay, M. (2015). Characterizing evidentiality. Linguistic Typology, 19(2), 201-237.

[13] Hall, E. T. (1976). Beyond Culture. Anchor Books, New York.