

Pragmatic Roles of Visuals in ELT Textbooks

Chie Kawashima

Tochigi Technical High School

Reference Data

Kawashima, C., (2021). Pragmatic roles of visuals in ELT textbooks. In P. Clements, R. Derrah, & P. Ferguson (Eds.), *Communities of teachers & learners*. JALT. <https://doi.org/10.37546/JALTPCP2020-36>

This study explored how visual elements are used in textbooks to convey information about pragmatics. The data analysed was visual information collected from beginner level ELT textbooks. The outcome of the study revealed: (a) 4-5 percent of the visuals provide contextual information associated with speech acts, (b) the majority of these visuals represent speaker-hearer power relations and social distance in relation to language use, and (c) these visuals tend to be associated with the speech acts particular to the language use of the speaker-hearer relationship depicted in the visual. In addition, the weakness of visuals as contextual information, which provides the viewers few opportunities to be aware of indirectness or politeness to be used in certain situations, was discussed accordingly. Some practical suggestions were made as to how teachers might supplement contextual information provided with visuals and compensate for their insufficiencies.

本論においては、語用論に関する情報伝達に関して語学学習教材の中で紹介されている視覚的要素の有用性の探査を、初級レベル学習者向けの英語教材から収集した視覚情報の分析により行った。その結果、教材で紹介されている発話行為使用における話し手と聞き手の関係が描写されている視覚要素の割合、話し手と聞き手の力関係ならびに社会的距離、またそれに伴う発話行為の種類が明らかになった。したがって、それぞれの話し手と聞き手の関係において特徴的な言語使用が検証され、同時に丁寧表現や間接的表現の使用場面の欠如といった問題点も指摘された。最後に、教材に含まれている視覚情報に関連する言語使用情報の不十分さを補うための教師作成による補助教材の実用性も示唆された。

Appropriate use of speech acts in interpersonal interactions is a crucial aspect of pragmatic competence. In many cases, the speech acts presented in ELT textbooks are introduced and practiced without any speaker-hearer information (Alcon-Safont,

2001; Boxer & Pickering, 1995; Nguyen, 2011). For novice language learners, visual materials are beneficial to comprehend the texts (Hadley, 2001) and beginner-level ELT textbooks are full of visuals such as photos and graphical illustrations. This study looks at visuals associated with speech acts to determine how each of them conveys contextual information of speech acts to learners by categorising visuals into four types according to speaker-hearer power relations and social distance. Based on the findings, how these visuals can help learners to select language use is discussed.

Literature Review

Pragmatic Role of Visual Images

Visuals help explain concepts and words thoroughly and precisely. For example, images in advertisements convey attractive dreams or potential future pleasures to viewers, and images appearing in newspapers show the critical moment of an incident or an accident (Van Leeuwen, 2008). Furthermore, Feng and O'Halloran (2012) explored how emotive meaning was represented in comic book drawings by analysing the features of emotions with facial expression semiotic resources including the shape of the mouth, wrinkles, cheeks and eyes, touch, and body orientations. As mentioned above, visual images may play meaningful roles in our everyday communication. They have various functions, such as stimulating the cognitive processes of viewers and inferential abilities for interpreting materials.

Visual Images in Language Textbooks

The purpose of presenting visuals is one of the crucial factors to determine whether or not the images can facilitate the textbook user's language development. Classification of the visual images in ELT textbooks as "for use" and "for decoration" has been made in a number of studies (Basal et al., 2016; Hill, 2003; Romney, 2012; Romney & Bell, 2012; Roohani & Sharifi, 2015). Although many of the textbooks' visuals were classified into the

'representational' functional category, Basal et al. (2016) pointed out that a still higher number of visuals with decorative functions were problematic. Both Romney and Bell (2012) and Basal et al. (2016) argued that the problem with decorative images is to attract the user's attention for aesthetic reasons with commercial purposes.

In addition, a study by Weninger and Kiss (2012), which examined the interplay of texts, images, and tasks in school EFL textbooks, found the materials inadequate for fostering intercultural competence and global citizenship as their primary goal of presenting the visual images was to extract denotational meanings. The following year, Kiss and Weninger's (2013) study of the interaction of texts and visuals in pedagogical materials noted the lack of tasks with visuals included in ELT resources. Their study adapted Tseng's (2002) triangle way of meaning-making process for interpreting culture, which suggests the role of cultural learning as developing the individual's ability to accept the diversity of other cultural forces. Kiss and Weninger (2013) noticed the significance of using visuals in ELT resources for classroom instruction. In turn, they suggested using text and visuals to generate a critical discussion at the beginning of a lesson based on the learners' spontaneous reactions, where the teacher assists them with "the process of recognition, exploration, and reflection" (Tseng, 2002, p. 16).

Aim of this Study and Research Questions

The motive for analysing visuals in this study originated in the observation that the speaker-hearer relationship regarding speech acts identified in the ELT textbooks was shown with visual images. For example, the visuals which depict the situation of service encounters are involved in the speech acts of *giving offers* and *requesting*, and conventionally indirect forms are used for the most part. As beginner-level ELT textbooks include a number of visual images, I have assumed that these visuals might provide the viewers with contextual information regarding the use of speech acts. Although there have been studies of the pragmatic roles of visual images that were not relevant to pedagogical issues, few studies, if any, have focused on the pragmatic use of visuals in pedagogical materials. In order to determine how the visual elements are used in the textbooks to convey information about pragmatics, the following three research questions were developed:

- RQ1. What percentage of visuals in each ELT textbook provides contextual information of speech acts?
- RQ2. What are the power relations and social distance of the speaker and hearer in each of the visuals associated with speech acts?

- RQ3. What types of speech acts or linguistic forms are involved in each category of the speaker and hearer's power relation and social distance?

Method

Textbook Selection

In order to compare two different types of ELT textbooks as to inclusion of pragmatic information, international ELT textbooks designed based on a situational syllabus and textbooks used in Japanese high schools, which are developed along with the grammar-based syllabus, were selected. To select the materials to be analysed, emails were sent to the Japan office of five major international publishers to enquire about their top-selling international ELT coursebooks at an equivalent level of the first year of upper secondary education (grade 10). Based on their replies, the most widely used materials to develop adult learners' four language skills identified by each publisher were selected (see Appendix A). In addition, seven ELT textbooks for first-year students used in Japanese high schools were also selected for analysis (see Appendix A). This textbook selection was also based on the information about the most widely used textbooks provided by the sales representatives from the major Japanese publishers. The elementary stage levels and the target users vary across the textbooks. *Progress*, *Treasure*, and *Crown* are mainly used in high level academic high schools. *Vista* and *Select* are designed for the curriculum implemented in vocational high schools. *Vivid* and *Quest* are used in all types of schools.

Selection of Visual Images

Before analysing the contextual information in each image, the total number of visual images presented in each textbook was counted, and the average number of visuals per page of each textbook was calculated. In this study, all the photos and illustrations were counted as visual images. Those with single frames were counted as one. A four-frame comic strip was counted as four.

Subsequently, the number of visual images which depict a speaker-hearer relationship in each textbook was determined to calculate an overall percentage of visuals that provided contextual information of speech acts. The images were judged to contain contextual information if they include narrative process of ideational meta-function based on Kress and van Leeuwen's (2006) social semiotic theory. Ideational meta-function refers to the representation of the objects, and narrative process represents participant actions, events, or processes of change. For example, the visual which depicts two human images at the cashier in a shop was judged to include narrative process of

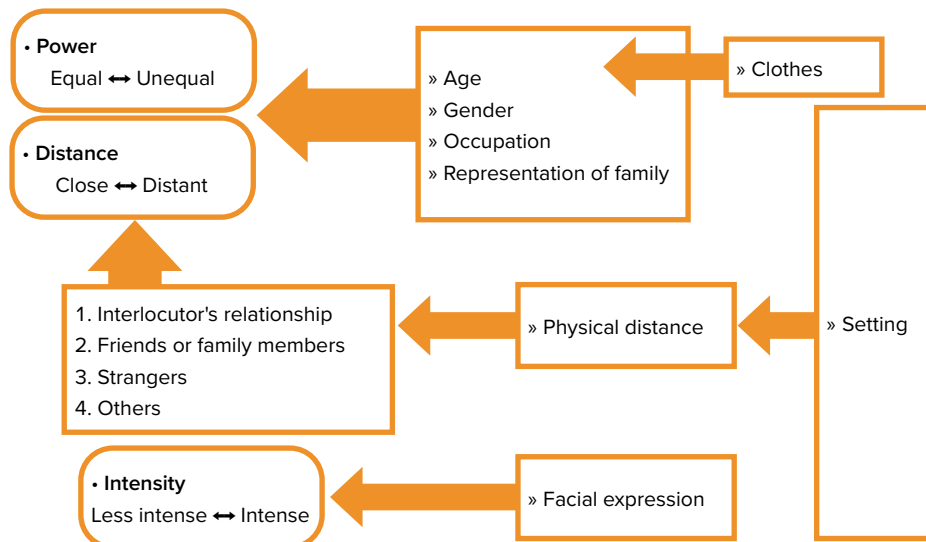
ideational meta-function and counted as the visual providing contextual information associated with the speech acts for service encounters.

Categorisation of Visual Images

To categorise the visual images presented in these ELT textbooks, two instruments were created as shown in Figure 1 and Figure 2 based on Kress and van Leeuwen's (2006) social semiotic theory and Ishihara and Cohen's (2010) use of instruments for field observations or for recording natural conversations.

According to van Leeuwen and Jewitt (2011), meaning is established through the syntactic relations among characters, locations, and things depicted in visuals with social semiotic visual analysis. Each element in Figure 1 provides information to determine the speaker and hearer's power relations, distance, and the intensity of the situation. Intensity refers to the imposition in a request or the severity in an apology as how great the request being made is or how oppressive the mistake or errors was, respectively (Ishihara & Cohen, 2010; Leech, 2014; Thomas, 1995).

Figure 1
Speaker-hearer Power Relations and Determinants



Whether the participants' power is equal or unequal was determined, referring to the categories of *power relations* summarised by Thomas (1995), as follows:

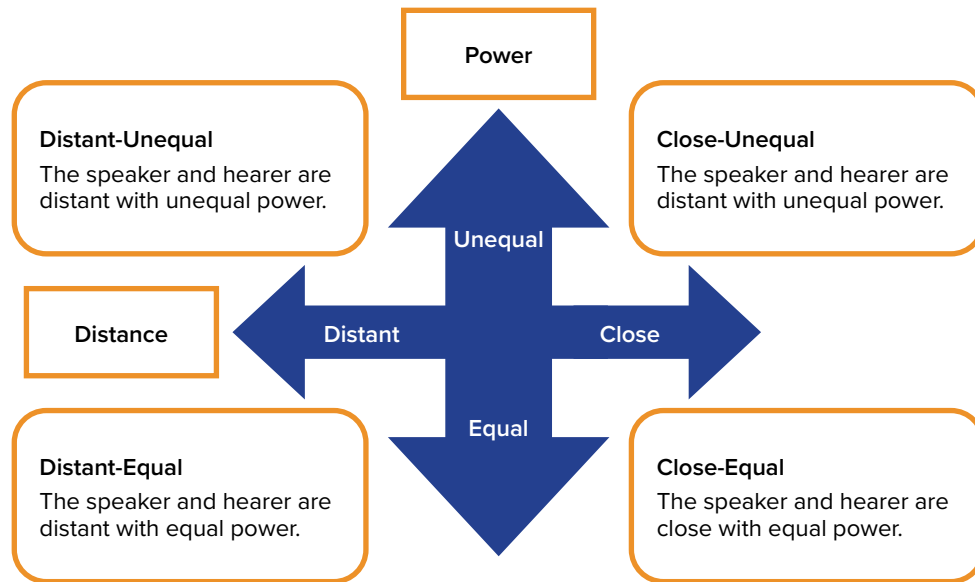
- a. Legitimate power (reward power and coercive power) – power is used by the speaker's virtue of role, age, or status such as employer-employee, parent-child, teacher-student.
- b. Referent power – someone who is admired/respected or can be a role model has power over another.
- c. Expert power – power is possessed by those with some special knowledge or expertise that the other person needs.

Additionally, the power possessed by someone with a specific right in a specific situation such as a police officer, a security guard and an immigration officer was also taken into consideration.

The classification of *distance* was made based on the variables as: a) social similarity/difference, b) frequency of contact, c) length of acquaintance, d) familiarity, or how well people know each other, and e) sense of like-mindedness (Spencer-Oatey, 1996). A number of contextual factors impacting the speaker and hearer's information helped to identify their power relations and social distance of the speaker and hearer. *Age, gender, occupation, and representation of family* are relevant contextual elements to determine speaker-hearer power relations and distance (Cohen & Ishihara, 2013; Ishihara & Cohen, 2010). These elements can also influence how socially and psychologically distant/close people feel to each other (Ishihara & Cohen, 2010; Spencer-Oatey, 1996). In addition, the participant's *clothes* can display cultural categorisations representing gender, institutional identity, ethnic identity, or relational identity (Van Leeuwen, 2008). When represented in a visual image, physical distance, or personal space such as public, social, personal, and intimate distance can also provide input on speaker-hearer relationship, although it is culturally specific (Cherry, 2018).

The visuals identified as contextual information of speech acts were categorised into four types according to speaker-hearer power relations and distance based on the diagram in Figure 2.

Figure 2
Categorisation of Visual Images According to Power Relations and Distance



The visuals categorised into *distant-unequal* included mainly doctor-patient, teacher-students, and so on. The visuals included in *close-unequal* depicted mostly parents and children. The speaker and hearer in the visuals categorised into *distant-equal* were mainly those in service encounters. The visuals which depict friends or family members were typically categorised into *close-equal*. In cases where the image showed the speaker and hearer's similarity but short length of acquaintance, their distance was determined by physical distance or the setting. Some visuals included the speaker and hearer who were not clearly enough depicted to be categorised into one of these four types, so *others* was added as a miscellaneous category.

Identification of Speech Acts

First, speech acts appearing in each textbook were identified using a comprehensive list of speech acts (see Appendix B). This list was created based on combining lists from the previous studies on the speech acts presented in ELT textbooks (Arai, 2005; Nguyen,

2011; Shimizu et al., 2007; Vellenga, 2004). Then, further investigation was undertaken to determine the number of speech act types involved in the visuals related to speaker-hearer relationship as categorised into four types in Figure 2. Linguistic forms involved in each of these speech acts were also examined to note whether there were any specific forms related to politeness, directness, or formality specific to the different power relations and distance of the speaker and hearer.

Results

Visual Images Providing Contextual Information in Each Textbook

The number of visual images which provide contextual information of speech acts in each textbook is displayed in Tables 1 and 2 along with the outcomes of a page-by-page analysis of those included in each textbook. Although the average number of visual images included in the set of international ELT textbooks is far more extensive than those used in Japanese high schools, the average percentage of the images which represent contextual information is only 4-5% and does not make any noticeable difference between these two sets of textbooks.

Table 1
Visuals Presented in the Set of International ELT Textbooks

Textbook	Interchange	Headway	Cutting	Global	Time
Number of pages	160	140	128	128	144
Average number of visuals per page	3.1	3.5	2.6	3.4	3.3
Total number of visuals	499	487	329	438	477
Number of visuals with contextual information (%)	35 (7%)	41 (8.4%)	18 (5.5%)	3 (0.6%)	15 (3.1%)

As a whole, a number of visuals which provide contextual information are present in each of international ELT textbooks. Only in *Global*, however, the number of visuals with contextual information is exceptionally limited although the total number of visuals is far larger than that of *Cutting*.

Table 2
Visuals Presented in the Set of Japanese High Schools Textbooks

Textbook	Progress	Treasure	Crown	Vista	Vivid	Select	Quest
Number of pages	243	215	191	135	160	103	138
Average number of visuals per page	1.1	0.5	1.0	1.9	1.0	1.8	1.0
Total number of visuals	260	100	184	262	162	189	138
Number of visuals with contextual information (%)	26 (10.0%)	0	2 (1.0%)	11 (4.2%)	1 (0.6%)	11 (5.8%)	9 (6.5%)

As for the set of textbooks used in Japanese high schools, the percentage of visuals with contextual information is uneven across the textbooks. In *Progress*, as many as 26 visuals, which account for 10 percent of the total number of visuals, represent contextual information while the number of visuals with contextual information is very limited in the remaining textbooks. In *Treasure*, these visuals are completely absent.

Speaker-Hearer Power Relations and Social Distance

Most of the visuals presented across the textbooks were categorised into *distant-equal* or *close-equal*, where speaker-hearer power relations are equal. As for the social distance, nearly half of the visuals included in the set of international ELT textbooks were categorised into *close-unequal* or *close-equal*. Likewise, 60 percent of the visuals in those used in Japanese high schools can be classified into *close-unequal* or *close-equal*. Therefore, most of the visuals included across the textbooks represent equal speaker-hearer power relations regardless of their social distance.

The visuals which depict the speaker and hearers at service encounters as well as those who are asking/giving directions were categorised into *distant-equal*. This type of visuals accounts for 41 percent of those with contextual information across the international ELT textbooks and 22 percent across the textbooks used in Japanese high schools. Finally, *close-equal* visuals, where friends, couples and siblings are depicted, make up over 40 percent across both sets of textbooks.

The visuals categorised into *distant-unequal* account for nine percent across international ELT textbooks and 18 percent across the textbooks used in Japanese high schools. As for detailed speaker-hearer relationships, teacher-student relationship is depicted in as many as 13 out of 21 visuals. Concurrently that of doctor-patient can be categorised into this type. Moreover, the visuals depicting an immigration officer and a passenger who is crossing the border at the passport control were also classified into *distant-unequal*, as well as the one with a security guard and a group of audience at the entrance of a football stadium. *Close-unequal* visuals, which account for four percent across the international ELT textbooks and 13 percent across the textbooks used in Japanese high schools, all depict family members like parent-child. Additionally, five visuals in *Headway* were impossible to categorise into one of these four types of power relations and social distance due to minimal speaker and hearer's information depicted and were categorised into *others*.

Linguistic Forms Associated Concerning Speaker-Hearer Power Relations and Distance

The range of speech acts involved in each type of visuals are shown in Tables 3 and 4. Although the total number of samples presented in the set of international ELT textbooks is larger, the range of speech acts in the visuals with contextual information in the Japanese high school textbooks is slightly more expansive than the international ELT textbooks, except for *distant-equal* relationships. As a whole, a wider range of speech acts is associated with the visuals in *distant-equal* and *close-equal*, while the visuals in *distant-unequal* and *close-unequal* involve a relatively narrow range of speech acts. Therefore, the range of speech acts associated with visuals might be tied up with the number of those categorized into each group.

Table 3

Speech Acts Expressed in Visuals in International ELT Textbooks

Type (Number of Visuals)	Speech Acts
Distant-Unequal (10)	Prohibiting, Requesting, Thanking, Greeting, Leave-taking, Giving orders, Expressing sympathy [7]
Close-Unequal (4)	Suggesting proposals, Asking for permission, Requesting, Refusing requesting, Greeting, Giving orders, Apologising, Making accusations [9]
Distant-Equal (46)	Giving advice, Giving offers, Refusing offers, Requesting, Refusing requests, Apologising, Complimenting, Responding to compliments, Thanking, Responding to thanking, Greeting, Leave-taking, Giving orders, Expressing opinions, Asking directions, Giving directions, Expressing sympathy [17]
Close-Equal (47)	Giving advice, Accepting advice, Giving offers, Accepting offers, Refusing offers, Suggesting proposals, Accepting proposals, Refusing proposals, Apologising, Responding to apology, Complimenting, Responding to compliments, Thanking, Greeting, Leave-taking, Giving orders, Agreeing, Assessing, Encouraging, Celebrating, Exclaiming, Expressing sympathy [21]
Others (5)	Apologising, Complimenting, Responding to compliments, Thanking, Leave-taking [5]

Note: The number of speech act types which involve visuals in each type of speaker-hearer relationship is shown in brackets.

Table 4

Speech Acts in Visuals in Japanese High School Textbooks

Type (Number of Visuals)	Speech Acts
Distant-Unequal (11)	Giving advice, Giving offers, Suggesting proposals, Asking for permission, Giving permission, Prohibiting, Requesting, Complimenting, Greeting, Leave-taking, Giving orders, Assessing [12]
Close-Unequal (8)	Giving advice, Asking for permission, Requesting, Thanking, Giving orders, Expressing opinions, Assessing, Exclaiming, Responding to good news [9]
Distant-Equal (13)	Giving advice, Accepting advice, Giving offers, Accepting offers, Requesting, Thanking, Responding to thanking, Greeting, Asking directions, Giving directions, Assessing [11]
Close-Equal (28)	Asking for advice, Giving advice, Suggesting proposals, Accepting proposals, Asking for permission, Giving permission, Requesting, Apologising, Complimenting, Responding to compliments, Thanking, Responding to thanking, Greeting, Leave-taking, Giving orders, Asking for opinions, Expressing opinions, Agreeing, Disagreeing, Assessing, Exclaiming, Expressing sympathy [22]

Note: The number of speech act types which involve visuals in each type of speaker-hearer relationship is shown in brackets

The categorised visuals tend to be associated with the speech acts specific to the language use of the speaker-hearer relationships depicted in these visuals. The speaker-hearer relationship in many of *distant-equal* visuals is relevant to service encounters. The use of the speech act of *requesting* with conventionally indirect forms, e.g., “Can you ...?” or “Can I...?” as a fixed expression in service encounters, was observed. In *Headway* the linguistic form of “I would like ...” is highlighted as a polite request. In *Cutting*, on the other hand, most of the visuals associated with *requesting* are presented with another fixed expression of “N, please.” Moreover, the visuals associated with the speech act of *giving orders* were mainly categorized into *distant-unequal*. *Imperatives* are mostly used

by a teacher as a part of classroom English with *distant-unequal* visuals. Likewise, the speech act of *giving orders* is typically employed by someone with a specific right, such as an immigration officer at the passport control and a security guard in a football stadium depicted in *distant-unequal* visuals.

Discussion

Although the page-by-page analysis of each textbook revealed inclusion of a large number of visual images, the percentage of those which are associated with speech acts is very small. Nevertheless, a wide range of speech acts is associated with these visuals across the textbooks. There is a pattern in involving speech acts in the visuals categorised into each type of speaker-hearer relationship.

The current study noted that the speaker and hearer's social role may have an impact on language use involved in the visuals categorised into *distant-equal* and *distant-unequal*. The speech act of *requesting* is typically present with *distant-equal* visuals in service encounters in these textbooks. Although this speech act is politeness sensitive (Leech, 2014), linguistic forms introduced along with these visuals are conventionally indirect or *imperatives + please* for the most part. Spencer-Ottay (1996) regards the role relationship used for the dyads in service encounters as equal since the service givers' right to receive payment and obligation to provide good service and service receivers' right to complain about poor service and obligation to pay are balanced out. As for language use, Gagne (2018) observed principal use of conventionally indirect request for product request in the authentic data. Therefore, direct or conventionally indirect language use may predominantly appear along with the visuals of *distant-equal* as far as the speaker and hearer's social role is concerned.

Indirect language use with *distant-unequal* visuals is entirely absent across the textbooks, and the language use with these visuals is typically direct. Although indirect use of language is governed by the factors represented in unequal power relations with social distance between the speaker and hearer along with rights and obligations (Thomas, 1995), only the language use of those with higher power is highlighted in these textbooks. As Taguchi and Rover (2017) describe, speakers' power is firmly controlled by the context and social roles rather than intrinsic to individuals. It may be necessary to be aware that the speaker has power over the hearer in a particular context while their power relation may be converted in another context with different social roles. In the case of the utterance of "Move this vehicle" produced by a police officer, this imperative is made in the capacity of a police officer although s/he might choose indirect language

use to say the same thing in his/her private capacity (Thomas, 1995). If the language use of those with lower power was also highlighted, indirect/polite language use might be introduced along with these visuals. In contrast, language use of the speaker and hearer in both *close-unequal* and *close-equal* visuals is not influenced by their social roles.

Overall, indirect use of language associated with these visuals is almost absent. Although it is preferable to be indirect, polite, or formal due to a number of factors in certain circumstances, almost all the linguistic forms of the speech acts associated with the visuals identified in this study are direct or conventionally indirect. Therefore, the contexts for language use are limited to those where indirect/polite language use is not required. Perhaps, the syntax structures of dialogues were simplified due to the necessity to match the language use to the learners' proficiency level, or keep the linguistic focus of the unit and important aspect of social interaction is neglected (Nguyen & Ishitobi, 2012).

Conclusion

Although only 4-5 percent of all the visuals included in the ELT textbooks were analysed in this study, these visuals are associated with a number of different speech act types. Most of these visuals were categorised into *distant-equal* and *close-equal*. The result of the categorisation of these visuals uncovered that language use employed in a specific situation is closely related to the speaker and hearer's social roles.

Although few visuals which could provide an opportunity to use indirect/polite forms were included in the textbooks, visual images may still help textbook users with selecting language use according to the circumstances and speaker-hearer relationship. As Kiss and Weninger (2013) noted the significance of visuals presented in ELT materials for classroom instruction and the potentials to enhance learners' cultural awareness through their critical discussion, visual images may also play a key role for pragmatic language instruction.

The activities to enhance the learners' cultural awareness using visual images proposed by Kiss and Weninger (2013) could be adapted to pragmatic input with visuals, where learners can discuss the language use in relation to speaker-hearer power relations and social distance. As Ishihara and Cohen (2010) suggested, an awareness-raising approach along with explicit instruction according to a list of tasks to analyse and identify language use both with linguistic and sociocultural focus may help learners notice the significance of selecting language use in contexts. Likewise, the necessity of raising awareness through discussion about language use according to speaker-hearer

relationship has been emphasised (Eslami & McLeod, 2010; Rinnert & Iwai, 2010; Yates & Springall, 2010). Therefore, visual images, as such, can be used for pragmatic input and practice raising learners' awareness of appropriate language use according to the context with a teacher's modification.

Bio Data

Chie Kawashima is an English teacher at Tochigi Prefectural Tochigi Technical High School. She is currently writing a doctoral thesis on pragmatic information presented in ELT materials. <kawashimachie@hotmail.com>

References

- Alcon, E., & Safont, P. (2001). Occurrence of exhortative speech acts in ELT materials and natural speech data: A focus on request, suggestion and advice realisation strategies. *Studies in English Language and Linguistics*, 3, 5-22.
- Arai, T. (2005). Teaching politeness in lower secondary schools in Japan: Text analysis of politeness. *Proceedings of the 10th conference of Pan-Pacific association of applied linguistics*, 15-30.
- Basal, A., Celen, K. M., Kaya, T., & Bogaz, S. N. (2016). An investigation into illustrations in English course books in a Turkish context. *International Electronic Journal of Elementary Education*, 8(3), 525-236. Retrieved from <https://www.iejee.com/index.php/IEJEE/article/view/129>
- Boxer, D., & Pickering, L. (1995). Problems in the presentation of speech acts in ELT materials: the case of complaints. *ELT Journal*, 49(1), 44-58. <https://doi.org/10.1093/elt/49.1.44>
- Cherry, K. (2018). Understanding body language and facial expressions Retrieved from <https://www.verywellmind.com/understand-body-language-and-facial-expressions-4147228>
- Cohen, A., & Ishihara, N. (2013). Pragmatics. In B. Tomlinson (Ed.), *Applied linguistics and materials development* (pp. 113-126). Bloomsbury.
- Eslami, Z. R., & McLeod, K. D. (2010). It's 8 o'clock in the morning. Are you watching television? Teaching indirect requests. In D. Tatsuki & N. R. Houck (Eds.), *Pragmatics: Teaching speech acts* (pp. 29-28). Teachers of English to Speakers of Other Languages.
- Feng, D. Z., & O'Halloran. (2012). Representing emotive meaning in visual images: A social semiotic approach. *Journal of Pragmatics*, 44, 2067-2084. <https://doi.org/10.1016/j.pragma.2012.10.003>
- Gagne, C. (2018). Indirectness and entitlement in product requests in British service encounters. *Journal of Pragmatics*, 133, 1-14. <https://doi.org/10.1016/j.pragma.2018.05.015>
- Hadley, A. O. (2001). *Teaching language in context*. Heinle & Heinle.
- Hill, D. A. (2003). The visual elements in EFL coursebooks. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 157-166). Continuum. <https://doi.org/10.5040/9781474211826.ch-010>
- Ishihara, N., & Cohen, A. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Pearson. <https://doi.org/10.4324/9781315833842>
- Kiss, T., & Weninger, C. (2013). A semiotic exploration of cultural potential in EFL textbooks. *Malaysian Journal of ELT Research*, 9(1), 19-28.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. Routledge. <https://doi.org/10.4324/9780203619728>
- Leech, G. (2014). *The pragmatics and politeness*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195341386.001.0001>
- Nguyen, M. T. T. (2011). Learning to communicate in a globalized world: To what extent do school textbooks facilitate the development of intercultural pragmatic competence? *RELC Journal*, 42(1), 17-30. <https://doi.org/10.1177/0033688210390265>
- Nguyen, H. T., & Ishitobi, N. (2012). Ordering fast food: Service encounters in real-life interaction and in textbook dialogues. *JALT Journal*, 34(2), 151-185 <https://doi.org/10.37546/jaltj34.2-2>.
- Rinnert, C., & Iwai, C. (2010). I want you to help me: Learning to soften English requests. In D. Tatsuki & N. R. Houck (Eds.), *Pragmatics: Teaching speech acts* (pp. 29-46). Teachers of English to Speakers of Other Languages.
- Romney, C. (2012). Images in ELT textbooks: Are they just decoration? In A. Stewart & N. Sonoda (Eds.), *JALT 2011 Conference Proceedings* (pp. 392-398). JALT. Retrieved from <https://jalt-publications.org/files/pdf-article/jalt2011-038.pdf>
- Romney, C., & Bell, L. (2012). The role of graphics in business English. In R. Chartrand & E. Skier (Eds.), *The 2011 Pan-SIG Conference Proceedings* (pp. 210-219). JALT. Retrieved from <https://pansig.org/publications/2011/pansig2011proceedings.pdf>
- Roohani, A., & Sharifi, M. (2015). Evaluating visual elements in two EFL textbooks. *Indonesian Journal of Applied Linguistics*, 4(2), 68-77. <https://doi.org/10.17509/ijal.v4i2.684>
- Spencer-Ottay, H. (1996). Reconsidering power and distance. *Journal of Pragmatics*, 26, 1-24. [https://doi.org/10.1016/0378-2166\(95\)00047-x](https://doi.org/10.1016/0378-2166(95)00047-x)
- Shimizu, T., Fukasawa, E., & Yonekura, S. (2007). Introductions and practices of speech acts in Oral Communication 1 textbooks: From a viewpoint of interlanguage pragmatics. *Sophia Linguistica*, 55, 143-163.
- Taguchi, N., & Rover, C. (2017). *Second language pragmatics*. Oxford University Press.
- Thomas, J. (1995). *Meaning in interaction: An introduction to pragmatics*. Longman
- Tseng, Y. H. (2002). A lesson in culture. *ELT Journal*, 56(1), 11-21. <https://doi.org/10.1093/elt/56.1.11>
- van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford University Press.

van Leeuwen, T., & Jewitt, C. (2011). Introduction. In T. Van Leeuwen & C. Jewitt (Eds.), *The handbook of visual analysis* (pp. 1-9). SAGE Publication.

Vellenga, H. (2004). Learning pragmatics from ESL & EFL textbooks: How likely? Retrieved from <http://tesl-ej.org/ej30/a3.html>

Weninger, C., & Kiss, T. (2012). Culture in English as a foreign language (EFL) textbooks: A semiotic approach. *TESOL Quarterly*, 47(4), 694-716. <https://doi.org/10.1002/tesq.87>

Yates, L., & Springall, J. (2010). Soften up! Successful requests in the workplace. In D. Tatsuki & N. R. Houck (Eds.), *Pragmatics: Teaching speech acts* (pp. 67-86). Teachers of English to Speakers of Other Languages.

Appendix A

Textbooks Used for This Study

International ELT Textbooks

- Interchange Intro (Cambridge University Press 2013)
- American Headway Starter (Oxford University Press 2010)
- Cutting Edge Starter (Pearson Education 2012)
- Global Beginner Course Book (MacMillan 2010)
- Time Zones 1 (Cengage Learning 2010)

Textbooks Used in Japanese High Schools

- Progress IN ENGLISH 21 BOOK 3 (Edec 2012)
- New Treasure English Series Stage 3 (Z-kai 2003)
- Crown English Communication I (Sanseido 2013)
- Vista English Communication I (Sanseido 2013)
- Vivid English Communication I (Daiichi Gakushusha 2013)
- Select English Expression I (Sanseido 2013)
- Vision Quest English Expression I (Keirinkan 2013)

Appendix B

List of Speech Acts

Speech Act Categories	Speech Acts
1. Advice	Asking for advice
	Giving advice
	Accepting advice
2. Offers	Giving offers
	Accepting offers
	Refusing offers
3. Proposals	Suggesting proposals
	Accepting proposals
	Refusing proposals
4. Permission	Asking for permissions
	Giving permissions
5. Prohibition	Prohibiting
6. Requests	Requesting
	Refusing requesting
7. Apology	Apologising
	Responding to apology
8. Compliments	Complimenting
	Responding to compliments
9. Accusation	Giving accusations
10. Thanks	Thanking
	Responding to thanking
11. Greetings	Greeting
	Leave-taking
12. Orders	Giving orders

Speech Act Categories	Speech Acts
13. Opinions	Asking for opinions
	Expressing opinions
	Agreeing
	Disagreeing
	Agreeing and disagreeing
14. Directions	Asking directions
	Giving directions
15. Assessment	Assessing
16. Celebration	Celebrating
17. Encouragement	Encouraging
18. Exclamation	Exclaiming
19. Sympathy	Expressing sympathy
20. Responses to good news	Responding to good news

Note: This list was created based on the studies of Arai (2005), Nguyen (2011), Shimizu et al. (2007) and Vellenga (2004).