

Research note

Which Preposition Do You Prefer, *Cut In* or *Cut Into*?: A Corpus-Based Analysis of *Cut In* Versus *Cut Into* Constructions

FUJIWARA Takafumi and IIJIMA Hisanori

英語話者は *Cut in* と *Cut into* のどちらを好むのか
—コーパスを用いた *Cut in* と *Cut into* 構文の使用実態に関する研究—

藤原 隆史・飯島 尚憲

Summary

The current study analyzes the phrase “cut *in/into* pieces” from cognitive semantics and pragmatics perspectives. The literature review revealed that the usage of *in* indicating resultative states can be traced to the usage in Old English. To check the actual usage of the phrase, the current study conducted the investigation using the COHA. The results show that the concerned usage of *in* is being replaced by *into*, which can denote a resultative state of Trajectors without any pragmatic or encyclopedic information.

Keywords

cut in/into, preposition in, preposition into, corpus linguistics, COHA

Contents

I. Introduction

II. Previous studies

III. Methodology

IV. Results and discussion

V. Conclusion

Acknowledgments

Notes

References

- e. cut a pie in four pieces (Wisdom)⁶⁾
 f. We have some cut fish that has been cut
 in pieces. (COCA)⁷⁾

In (2a, b), the LMs could not be regarded as a container. Rather, the TRs will form the LMs, namely, the LM “a circle” consists of the TR “chairs” in (2a), and the LM “line” consists of the TR (people including “you” who are lining up/queueing). Similarly, in (2c, d), the LMs have no property of the container. Eguchi (2022) indicates that in (2c, d), the LMs denote the “resultative states” of the TRs “the meat” and “the paper” after the subjects “she” and “he” performing the action expressed by the verb “cut”. Likewise, the LMs in (2e, f) denote resultative states of the TRs “a pie” and “fish”, which are divided into pieces in either case of the active or passive voice. Be it noted that the verb “cut” used in (2c-f) is called a “resultative verb” and it intrinsically entails the result of the dividing action as a semantic concept. This characteristic of the resultative verb enables not only *into* but also *in* to be used in a sentence like (2a-f) that contains an action and its result.

Regarding the relationship between *in* and *into*, Eguchi (2022) points out that though *in* and *into* can be used interchangeably under certain conditions, the acceptability of using *in* depends on pragmatic elements, i.e., the property of the verb and/or the LMs which appear in the same sentence. For instance, as mentioned in (2e, f), *in* can be used to indicate the final state of the TRs, whereas, in the following sentences, the contexts or the pragmatic conditions tend not to or do not sanction the use of *in*.

- (3) a. She cut the paper $\{[?]in/into\}$ a circle/
 triangle.
 b. John pounded the metal $\{*in/into\}$
 pieces^{Note2}. (Eguchi, 2022)⁸⁾

In (3a), according to Eguchi (2022), even though the acceptability goes down, *in* can be used when

the verb is resultative, and the *in*-phrase serves as a “further specifier” of the result of the action. In other words, the verb “cut” in this sentence (ones in 2c-f as well) performs a semantic role of coding the result as well as the action. On the other hand, as in (3b), if the verb is not resultative (i.e., the verb does not entail a result), *in* is not sanctioned for coding the result of the action. That is, *in* cannot be used to mediate a cause-effect relationship between the action and the final state (divided pieces). In the case of (3b), the verb “pound” denotes just a simple action of hitting the metal which does not entail any result^{Note3} and consequently, another language device (*into* in this case) must take the semantic role of coding the result. Therefore, in (3b) one cannot use *in* instead of *into*.

As observed above, *in* can denote not only container-containee relations but also resultative states under certain conditions, otherwise, *in* cannot be used as a marker of resultative states. In the next section, preposition *into* which has similar semantic and syntactic properties will be mentioned in detail, especially why *into* can mediate the resultative states even if the verb is not resultative as in (3b).

2. Semantics of *into*

This section will begin with a simple observation of the preposition *into* in a dictionary. Oxford Advanced Learner’s Dictionary Online (OALD)⁹⁾ defines the preposition *into* as follows^{Note4}.

- (4) a. to a position in or inside sth: *Come into the house.*
 b. to a point at which you hit sb/sth: *The truck crashed into a parked car.*
 c. used to show a change in state: *The fruit can be made into jam.*
 d. used to show the result of an action: *He was shocked into a confession of guilt.*

The definition of *into* in OALD refers to not only the spatial relationship between objects but also the resultative states of TRs. In other words, *into* can profile the final position/state of the TR thereby connotating the path of the TR. This means that prepositions *in* and *into* differ from each other in the point of denoting the change of location, which entails the path of the TR at the same time. This difference can be confirmed by the following examples. In (5a), the event is occurring *within* the LM and contains no movement of the TR along a path, whereas in (5b), *into* implies the transfer of the TR from the outside of the LM to the inside of it^{Note5}.

(5) a. He ran in the room [the running is within the confines of the room]

b. He ran into the room [the TR begins outside and oriented towards the room, and runs, such that the TR comes to be located within the room]

(Tyler and Evans, 2003, p.199)¹¹

Tyler and Evans (2003) explain that the difference between (5a) and (5b) derives from the difference between the dative case and the accusative case used after preposition *in* in Old English (OE) period. At that time, they note, the preposition *in* mediated both static location and dynamic orientation of TRs *in/into* LMs by putting the dative case and the accusative case respectively in the slot of the complement of the preposition. By Middle English (ME) period, the latter use of *in* came to be replaced by the preposition *into* after omitting the dative and the accusative distinction. Therefore, the observation in (2), i.e., *in* has the container property and can signify the resultative state of TRs as well, could be described/accounted for by looking at the diachronic change of the meanings of *in* and *into*. This diachronic point of view will be mentioned again in the next section, referring to the OED. Before doing this, let us look at Eguchi (2022)'s

account on *into* and its semantic/pragmatic properties.

As mentioned in the previous section, Eguchi (2022) refers to the difference between *in* and *into* by stating the different nature of verbs in the sentence and the pragmatic settings related to resultative states. In (6), he additionally points out that the acceptability of the sentences could be concerned with the semantic elements of the verb and the pragmatic elements of the sentence.

(6) a. John broke the stick {in/into} pieces.

b. John broke the egg {[?in/into} the bowl.

c. He tore the paper {[?in/into} pieces.

d. He tore the shirt {[*in/into} bandages.

(Eguchi, 2022, p.253)¹²

In (6a), the verb “broke” is a resultative verb and it entails the intrinsic meaning of causing a state change, namely division or destruction. Eguchi (2022) clarifies that, in this context, the verb itself semantically entails the meaning of destruction and connotes the final state of the TR “the stick”, meaning that the action caused the TR to be divided. In other words, the division or destruction of the TR is semantically coded by the verb and *in* itself performs as a further specifier of the final state and does not have to entail the state change of the TR. This is because, Eguchi (2022) explains, that the action of breaking and the result of being divided are the same events, i.e., if one breaks something, the object inevitably will be divided^{Note6}. Eguchi (2022) remarks that though the acceptability drops, the same explanation can be applied to (6c). In (6b), on the other hand, there are two separate events, namely, the action of breaking and the transfer of the object. Between the two events, an interval exists, meaning that the first event of breaking does not entail the transfer of the egg. Hence, since *in per se* cannot mediate the transfer of the TR, not a semantic but a pragmatic motivation is required to indicate the transfer and the final position of the egg. Speakers

of English and others sharing the same culture may share encyclopedic knowledge that some *mashed* egg is mixed with some mayonnaise in a *bowl* to make sandwiches. Eguchi (2022) notes that by using this kind of knowledge (pragmatic information), speakers of English may accept the use of *in* in (6b)^{Note7}. In addition to this, according to Eguchi (2022), in (6d), using torn pieces of the shirt for bandages is a specific use, which is pragmatic information and, in such cases, *into* is the first choice. To sum up, *in* can be sanctioned when the verb is resultative, or the event has rich pragmatic/encyclopedic information. Let us leave the synchronic aspect of *in* and *into* and turn to diachronic arguments.

3. Oxford English Dictionary

From the consideration of Tyler and Evans (2003) in the previous sections, it has been confirmed that not only synchronic but also diachronic observation should be considered. In this section, by consulting the OED, the diachronic change of the meanings of *in* and *into* will be examined. The following (7) and (8) are excerpted descriptions of the relevant parts of *in* and *into* in the OED¹³.

(7) The description of *in* in the OED

The preposition *in* expresses the relation of inclusion, situation, position, existence, or action, within limits of space, time, condition, circumstances, etc.; also (like Latin *in*) from the earliest times it expresses motion or direction from a point outside to one within limits; these two senses being distinguished (as with *on prep.*) by the case of the word affected, the former taking, in Old English, the dative (rarely the instrumental) for earlier locative, the latter the accusative or case of motion towards. However, the Old English perspective was rather different from the modern, and the choice of case can sometimes seem counter-intuitive (compare e.g. use of the

accusative with verbs of striking: see quot. OE at sense Li.8a). In practice, even allowing for this, the distinction between the functions of dative and accusative was not always clear-cut, with many senses evidenced which take both cases. In Middle English the distinction of case disappeared, but *in* continued in both uses, in the sense of motion or direction gradually coming to be superseded in ordinary use by the compound *into prep.* . . .

(underlines ours)

(8) The description of *into* in the OED

In Old English the noun following *into* is usually in the dative case, although use with a noun in the accusative is also occasionally found; compare the typical behaviour of Old English *tō* described at *to prep.*

In early use the collocation partly takes the place of uses of the simple preposition *in* followed by a noun in the accusative case (or case of direction), as distinct from uses of *in* followed by a noun in the dative case, expressing location. This probably partly reflects the gradual process of loss of distinct case endings in English. However, uses of the simple preposition *in* expressing movement into a location have remained current alongside use of *into* (see in *prep.* IV.33), and conversely *into* is also found (in later use only in Scots) expressing location or position (see A.III.22). . . .

As clearly stated in (7) and (8), *in* mediated both static location and dynamic orientation of TRs in the past. Therefore, peripheral usages of *in* such as (2) could be vestiges of the old usages in OE and the synchronic observations of *in* and *into* should be traced to a relatively old usage. This drives us to the question of how the prepositions have been used in the history of the English language so far. That is, the frequency, pragmatic settings, syntactic rules, etc. of the prepositions in

a corpus should be investigated. By doing this, it is expected that the diachronic change of *in* and *into* and a habitat segregation of them would be revealed.

4. Summary and research question

This chapter has looked at the semantics of *in* and *into* in terms not only of semantic/pragmatic but also of synchronic/diachronic point of view. From the observation above, when the verb is resultative or the context is rich in pragmatic/encyclopedic information, *in* can be used to indicate the final state of TRs (as in “cut *in* pieces”). Furthermore, it has been confirmed from the observation in the OED that this usage of *in* could be a trace of the usage of *in* with the accusative which indicated the final state of TRs in OE.

Here the question arises, namely, the question of what the frequency of occurrence of the prepositions in the phrases with a resultative verb in a certain corpus is. To put it differently, which preposition speakers of English prefer to use as a marker of the state change in a certain corpus is worth investigating. It is true that phrase such as “cut *in/into* pieces” has a resultative verb “cut”, and both *in* and *into* can mediate the final state of the TRs, but the frequency of occurrence of the prepositions in the phrase has not been investigated so far. Therefore, in the current study, the following research question will be set.

(9) Research question

What is the frequency of the occurrence of the prepositions *in* and *into* in the phrase “cut *in/into* n pieces”^{Note8} in the Corpus of Historical American English (COHA)?^{Note9}

As a hypothesis, the current study proposes that both structures are in the course of semantic habitat segregation where “cut *in* pieces” is fading

away, whereas “cut *into* pieces” is establishing its semantic position.

III. Methodology

As mentioned above, the current study adopted the methodology of corpus linguistics. Since the aim of the study is to reveal the chronological change of the concerned usage of *in* in a certain period in the history of the English language, the COHA was adopted. The COHA is a corpus that accumulates “more than 475 million words of text from the 1820s-2010s . . . and is balanced by genre decade by decade”¹⁴. The genres are novels, magazines, newspapers, and non-fiction. In the current study, the strings containing “cut *in/into* pieces” were retrieved from the COHA by using the “collocates” search. That is, the used phrases were “cut *in* + pieces” and “cut *into* + pieces” so that not only “cut *in/into* pieces” but also “cut *in/into* n pieces” such as “cut *in* four pieces” and “cut *into* small pieces” could be picked up. Moreover, in the collocates search, the word “pieces” was searched within nine words^{Note10} to the right of the phrase “cut *in/into*”.

To verify the significance of the difference between the frequencies of *in* and *into* during the period, the current study conducted a Chi-square test, dividing the period (sections) into two groups of the first period (the 1820s-1890s) and the second period (the 1900s-2010s). This division of the period is based on the periodization of the English language, namely, OE period (700-1100), ME period (1100-1500), Modern English (ModE) period (1500-1900), and Present Day English (PE) (1900-)¹⁵. From this periodization, the first period belongs to ModE and the second period belongs to PE. The Chi-square test compared the frequencies of occurrence of *in* and *into* in the first and the second periods. Regarding the formulation of hypotheses, the null hypothesis (H0) and the alternative hypothesis (H1) were formulated as follows: H0: There is no significant difference

between the occurrences of “cut *in n* pieces” and “cut *into n* pieces” in the periods, H1: There is a significant difference between the occurrences of “cut *in n* pieces” and “cut *into n* pieces” in the periods. In addition, the significance level (α) was set at 0.01.

IV. Results and discussion

Table 1 below shows the frequency of occurrence of the phrase “cut *in/into n* pieces”. As you can see, the total frequency of “cut *into n* pieces” (317) is higher than that of “cut *in n* pieces” (106). Before 1900, however, though there was a fluctuation in each decade, the total frequency of “cut *in n* pieces” (63) and that of “cut *into n* pieces” (66) were not so different from each other. Considering the proportions of occurrence in total, it can be said that “cut *in n* pieces” appeared more frequently before 1900. More specifically, in the 1820s, the 1850s, the 1860s, the 1880s, and the 1900s, the frequencies of “cut *in n* pieces” are higher than those of “cut *into n* pieces”. Furthermore, except in the 1830s, the frequencies of “cut *into n* pieces” are not so much higher than those of “cut *in n* pieces”. These observations could indicate that, in the first period (the 1820s-1890s), both phrases evenly occurred in the corpus, and occasionally “cut *in n* pieces” was rather a predominant choice to describe the situation of division of objects. Conversely, after the 1900s, the frequencies of “cut *in n* pieces” in each section are equal or lower to/than those of “cut *into n* pieces”. Moreover, during the 1900s-2010s, the total frequency of “cut *into n* pieces” (251) is much higher than that of “cut *in n* pieces” (43). Especially, after the 1980s, the frequency of “cut *into n* pieces” is increasing rapidly, occupying about 79% of the total. On the contrary, after the 1980s, the frequency of “cut *in n* pieces” tend to decrease.

To check if there is a significant difference between the frequencies of occurrence of *in* and

into in the first and the second periods, a Chi-square test was performed, and the following results were obtained. Tables 2, 3, and 4 show the crosstabulation table for Periods and Prepositions, the results of the Chi-square tests, and the effect size, respectively. As the Tables indicate, the relationship between the variables (Periods and Prepositions) was significant, $\chi^2(1, N = 423) = 55.880, p = .00, \phi = .363$. The effect size was moderate. Based on the above, the null hypothesis was rejected, and the alternative hypothesis was adopted. In addition to this, the residual analysis revealed that the frequency of *in* in the first period was significantly high, and the frequency of *into* in the second period was significantly high. This means that *in* was used in the first

Table 1. The frequency of occurrence

Section	Frequency	
	cut <i>in n</i> pieces	cut <i>into n</i> pieces
1820	5	2
1830	10	23
1840	3	5
1850	8	7
1860	10	6
1870	1	3
1880	25	18
1890	1	2
1900	12	4
1910	3	3
1920	3	5
1930	3	5
1940	2	3
1950	2	4
1960	0	1
1970	4	5
1980	3	9
1990	5	41
2000	4	84
2010	2	87
total	106	317

period more often than *into* was, whereas the frequency of *into* overtook that of *in* in the second period, connoting the shift from the use of *in* to *into* has been proceeding. Especially, after the 1980s, the frequency of occurrence of *into* has been increasing sharply, establishing its position as the preferred of the two options. From the observation above, it seems reasonable to suppose that, in the COHA, the usage of *in* mediating the

resultative state of TRs is being replaced by the usage of *into* mediating the same state after 1900. Therefore, the hypothesis mentioned in Chapter 2 can be valid.

V. Conclusion

The current study analyzed the semantic properties of prepositions *in* and *into*, utilizing the

Table 2. The Crosstabulation table for Periods and Prepositions

		Prepositions						Total	
		<i>in</i>			<i>into</i>				
		N	%	Adjusted Residual	N	%	Adjusted Residual		
Periods	1820s-1890s	63	59.4%	7.5	66	20.8%	-7.5	129	30.5%
	1900s-2010s	43	40.6%	-7.5	251	79.2%	7.5	294	69.5%
Total		106	100%		317	100%		423	100%

Table 3. The results of the Chi-square tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	55.880 ^a	1	.00		
Continuity Correction ^b	54.073	1	.00		
Likelihood Ratio	52.818	1	.00		
Fisher's Exact Test				.00	.00
N of Valid Cases	423				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 32.33.

b. Computed only for a 2x2 table

Table 4. Effect size

		Value	Approximate Significance
Nominal by Nominal	Phi	.363	.00
	Cramer's V	.363	.00
	Contingency Coefficient	.342	.00
N of Valid Cases		423	

cognitive semantics and pragmatics perspectives. The previous studies argue that some of the usages of *in* can denote not only the typical meaning of the container image schema and its metaphorically extended meanings but also some peripheral ones such as expressing resultative states of TRs. From the observations, when the verb in a sentence is resultative or the context has rich pragmatic information indicating a resultative state, *in* as well as *into* can mediate the resultative state of the LM. This has been confirmed by consulting the OED, i.e., in the past, *in* indicated resultative states by putting an accusative as LMs. This suggests that the concerned usage of *in* in PE can be a vestige of the usage in the past. However, the actual usage has not been investigated. To check this, the current study conducted a corpus-based analysis of the usages of *in* and *into*, especially the phrase “cut *in/into* *n* pieces”, by utilizing the COHA and a statistical analysis of a Chi-square test. The results show that the frequency of “cut *in* *n* pieces” was significantly higher than that of “cut *into* *n* pieces” in the 1820s-1890s (ModE), while, the frequency of “cut *into* *n* pieces” has been increasing with significance in the 1900s-2010s (PE), meaning the shift from *in* to *into* is proceeding. It should be noted that since the current study analyzed one specific phrase “cut *in/into* *n* pieces”, there is room for further investigation into the resultative usage of *in/into*. For instance, phrases such as “cut *in/into* slices”, “break *in/into* pieces”, or more schematically, the construction of “(resultative) verb + *in/into* + resultative states” should be worth investigating. Moreover, pragmatic settings, syntactic rules, etc., which were left untouched in the current study, should also be research issues. These issues will be addressed in the future.

Acknowledgments

This study was supported in part by JSPS KAKEN Grants No. JP22K00746 and No. JP20K00841. The authors express our deep

gratitude. The authors also wish to thank Professor Yasuhide Numata at Hokkaido University for his helpful comments on the statistical analysis. But responsibility for the text (with any surviving errors/typos) rests entirely upon the authors.

Notes

Note1 In the realm of cognitive semantics, Landmark (LM) refers to an entity that is not salient and serves as a reference point or background and Trajector (TR) refers to an entity that is more salient than LM and likely to move or perform some other acts.

Note2 Eguchi (2022) explains that the [?] stands for the lower acceptability in the results of the Google search. Add to this, the [*] expresses that the sentence is ungrammatical.

Note3 Since the verb “pound” does not always entail the result of dividing an object into pieces, it cannot be used to express the final state of the object. In other words, even if one pounds an object with a hammer, the object does not necessarily get crushed, stretched, or smashed.

Note4 In (4), the usages of *time* are omitted. This is because the current study focuses only on the spatial usages and the usage of changing the state of something.

Note5 Kageyama (1997)¹⁰ also assumes that the Lexical Conceptual Structure (LCS) of *into* is FROM-ENTRANCE-of-X TO-INSIDE-of-X.

Note6 In terms of aspect, the construction such as (6a) is a resultative one and aspect is telic, meaning the telicity is achieved only by the verb “break” *per se*.

Note7 Eguchi (2022) notes that, in (6b), not only *into* but also *in* is licensed to appear because the LM is a place. However, he admits that a further survey should be done to confirm this hypothesis.

Note8 The phrase “cut *in/into* *n* pieces” contains examples such as “cut *in/into* *four* pieces” or “cut *in/into* *small* pieces” as well as “cut *in/into* pieces”. The details are explained in Chapter 3.

Note9 It must be noted that as it is practically impossible to investigate all the corpora in the world, it could be reasonable to select one suitable corpus to collect data. That is why the current study has decided to utilize the COHA (to be explained in detail in Chapter 3).

Note10 The current study searched within nine words to the right of the target phrase because the default setting of the maximum range in the COHA is nine words to the left/right of the target word/phrase.

References

- 1) Lee D, *Cognitive Linguistics An Introduction*, Oxford University Press (2001).
- 2) Tyler A and Evans V, *The Semantics of English Prepositions Spatial Scenes, Embodied Meaning, and Cognition*, Cambridge University Press (2003), p.196.

Meaning, and Cognition, Cambridge University Press (2003).

- 3) Lee D, *Cognitive Linguistics An Introduction*, Oxford University Press (2001), p.19.
- 4) Tyler A and Evans V, *The Semantics of English Prepositions Spatial Scenes, Embodied Meaning, and Cognition*, Cambridge University Press (2003), p.196.
- 5) Eguchi T, “Henka Kekka-wo Arawasu in/into Zenchishiku-no Koutai nitsuite [A Study on the Phenomenon of Alternation of in/into which Denotes Variation Results], *In Chubu Eibungaku 14*, (2022), p.252.
- 6) Sanseidou, *Wisdom English-Japanese Dictionary 3rd edition*, (2013).
- 7) Davies M, *Corpus of Contemporary American English (COCA)*, <https://www.english-corpora.org/coca/> (2008-). (Retrieved on 2024.1.10).
- 8) Eguchi T, “Henka Kekka-wo Arawasu in/into Zenchishiku-no Koutai nitsuite [A Study on the Phenomenon of Alternation of in/into which Denotes Variation Results], *In Chubu Eibungaku 14*, (2022), p.252.
- 9) Oxford University Press, *Oxford Advanced Learner’s Dictionary Online*, <https://www.oxfordlearnersdictionaries.com/> (2023). (Retrieved on 2024.2.9).
- 10) Kageyama T, “Idou-to Kyori: deep into the forest toiu Hyougen-no Imikouzou [Movement and Distance: A Semantic Structure of the Expression “deep into the forest”], *Zinbun Ronkyu 47*(1), (1997), p.110.
- 11) Tyler A and Evans V, *The Semantics of English Prepositions Spatial Scenes, Embodied Meaning, and Cognition*, Cambridge University Press (2003), p.199.
- 12) Eguchi T, “Henka Kekka-wo Arawasu in/into Zenchishiku-no Koutai nitsuite [A Study on the Phenomenon of Alternation of in/into which Denotes Variation Results], *In Chubu Eibungaku 14*, (2022), p.253.
- 13) Oxford University Press, *The Oxford English Dictionary Online*, <https://www.oed.com/> (2000-). (Retrieved on 2024.2.9).
- 14) Davies M, *Corpus of Historical American English (COHA)*, <https://www.english-corpora.org/coha/> (2008-). (Retrieved on 2024.1.18).
- 15) Hashimoto I, *Eigo-shi Nyuumon [Introduction to the History of the English Language]*, Keio University Press, (2013[2005]), pp.64-65.