A Quantitative View of Short Utterances in Daily Conversation: A Case Study of *That's right*, *That's true* and *That's correct*

Yanjiao Li
Department of Linguistics
and Translation
City University of Hong
Kong
Hong Kong SAR
`yanjiaoli2-c`
`@my.cityu.edu.hk`

Alex C. Fang
Department of Linguistics
and Translation
City University of Hong
Kong
Hong Kong SAR
`acfang@cityu.edu.hk`

Jing Cao
School of Foreign Languages
Zhongnan University of
Economics and Law
Wuhan, P. R. China
`cecilia_cao`
`@znufe.edu.cn`

Abstract

Short utterances serve a multitude of different communicative functions in interactive speech and have attracted due attention in recent research in dialogue acts. This paper presents a quantitative description of three short utterances i.e. *that's right*, *that's true*, *that's correct* and their variations based on the Switchboard Dialogue Act Corpus. Particularly, it offers an overview to account for how they are deployed by native speakers in daily conversation. At the same time, it attempts to provide a comparative account of *that's right* and *that's true*, showing that while almost 75% of them are mutually exchangeable, they nonetheless exhibit preferences in interactive speech. This insight is expected to form a useful approach towards automatic dialogue act tagging.

1 Introduction

Dialogue act (DA), defined as "communicative activity of a dialogue participant, interpreted as having a certain communicative function and semantic content" (ISO 24617-2, 2012: 2), plays a key role in the interpretation of the communicative behaviour of dialogue participants and offer valuable insight into the design of human-machine dialogue system (Bunt et al. 2010). With the goal of facilitating automatic DA tagging, this paper describes a corpus-based investigation into *that's right, that's true, that's correct* and their variations in the Switchboard Dialogue Act (SWBD) Corpus, in order to answer questions about the communicative functions they mainly perform in daily conversation. These utterances deserve our particular attention in research considering that, like other brief responses (e.g. *Oh*, *Uh huh*, *Mm*, *Okay*), they serve as important feedback to the main speaker and they usually occur as overlapping speech. They are particularly problematic to interpret because they demonstrate a drastically different functional or pragmatic meaning from the semantic meaning of the component tokens. Consider Example 1.

Example 1
sd   B.54 utt1: -- {C and } I like that because [it's a, + it's] real easy to, {F uh, } follow for her, /
sd   B.54 utt2: {D you know, } {F uh, } {D gosh, } [if, + if] I read straight out of the Bible to her she'd <laughter> never understand any of it. /

sd   A.55 utt1: {D Well, } it's hard for me. /

ba   B.56 utt1: <u>That's right <laughter>. /</u>
                                           sw_0263_2226.utt

This is one excerpt retrieved from the targeted corpus, which will be further illustrated in section 2. A and B, two speakers, are talking about books and literature, where B is describing one of her daughter's book, "real easy to follow". The last utterance

that's right can be interpreted as serving both assessment/appreciation and agreement functions. The speaker B considers that what has been stated by A is right, not false, in which *right* is used as the evaluative adjective. Also, he implies his agreement with the interlocutor where *that's right* is used as a whole. Therefore, on the one hand, the semantic meaning of *that's right* makes it much closer to personal judgments and assessments, that is, the opinion is "right, not false". On the other hand, it is often used as a whole, indicating speaker's agreement, which goes beyond lexical meanings.

However, past studies rarely specify various usage for *that's right*, *that's true* and *that's correct* in a systematic fashion, and just sporadically describe one or two cases to illustrate one or two facets for them, without capturing a full picture of how they are used with empirical evidence. To be more exact, for the studies that do discuss usage for *that's right*, Gardner (2001) believes that *that's right* is exactly the same as *right* when responding to a preceding question, the synonym for "*that's correct*". This point has been further elaborated in that *right* is deemed "a truncated version of *that's right*" when acting as "an epistemic confirmation token", "in a sense close to one of its dictionary meanings, namely '*correct*'" (Gardner, 2004: 4). The studies indicate that *that's right*, *right*, *that's correct*, and *correct* are similar and can be alternatively used as the confirmation token oriented to a prior question. At this regard, however, Stenström (1987: 104) asserts that *that's right* is much stronger than *right* in degree of emphasis and involvement when severing as a response move to the same type initiating move. In addition, when responding to a previous declarative, *that's right* has been considered to realize the functions of seeking confirmation (Tui, 1994), showing agreement (Stenström, 1987; Tui, 1994; Gardner, 2001) as well as making assessments (Tao, 2003). Therefore, *that's right* has been considered to indicate a wide variety of intentions in interaction. With regard to *that's true*, it has received little attention, and only McCarthy (2003) makes brief description that as a syntactically independent token, *true* seems to prefer the clausal option (*that's true*) to independent occurrence (*true*). In terms of *that's correct*, it has been left largely unexamined and unspecified regarding the usage.

Considered semantic meanings of the three short utterances (i.e. *that's right*, *that's true* and *that's correct*), they largely embody in their key words *right*, *true* and *correct*. As is shown in dictionaries, the three words have similar lexical meanings and are often used to paraphrase each other. For instance, *Longman Dictionary of Contemporary English* (2009, fifth edition) defines them as follows:

Correct: having no mistakes; right (p.379)
Right: true/correct (p.1504)
True: not false, based on facts and not imagined or invented (p.1891-1892)

Thus, this paper aims to bring together the disparate findings on the uses of the three short utterances as well as their variations, attempting to depict an overview of them: how they are deployed by native speakers in daily conversation. At the same time, a comparative view has been concentrated on *that's right/true*, to seek to the circumstance in which they are mutually exchangeable and in which they are distinct. In this way, it is expected to form a useful approach towards automatic detection of DAs.

This paper is structured as follows. Section 2 briefly introduces the SWBD DA corpus, then section 3 presents how the data has been processed before statistical analysis. Section 4 is related to general figures for the three short utterances and their variations, followed by a comparative study (section 5). Section 6 draws conclusions to this paper.

2   Corpus Resource

This study uses the Switchboard Dialogue Act Corpus[1], which comprises 1,155 transcribed telephone conversations, totaling in 223,606 utterances or 1.5 million word tokens (Fang et al., 2011). In this corpus, the segmented unit for utterances is defined as "slash-unit", which can be complete or incomplete, ranging from "a sentence" to "a smaller unit" (Meteer et al., 1995: 16). Moreover, all these segmented utterances have been annotated with DA information, such as "aa" (*accept*), "ba" (*assessment/appreciation*), etc., to denote functions of particular utterances according to the SWBD-

---

[1] available online www.ldc.upenn.edu

DAMSL coding scheme (Jurafsky et al., 1997). Consider Example 2[2].

Example 2
sv   A.9 utt12: any jury's not going to disregard the evidence, {D you know } <laughter>. /

aa   B.10 utt1: {F Uh, } that's true. /
                                    sw_0142_2145.utt

As can be seen, the first utterance has been coded with "sv", a DA tag for *statement-opinion*, while the second one has been labeled as "aa", a code for *accept*. In the current study, investigation of various functions will be conducted based on the DA tags which have been coded for each utterance.

3   Data Pre-processing

For the benefit of the current work, *that's right*, *that's true* and *that's correct*, and their variations are retrieved from the corpus accordingly. Variations in the current study are defined with a series of factors taken into account.

- Firstly, variations of the same token share the key words and present in similar patterns, for instance, *it's true*, *this is true* and *true* are all considered as variations of *that's true*, since they contain the same key word *true* with similar patterns. Consequently, the whole utterances have similar semantic meanings.
- Secondly, cases (e.g. *it's true*) embedded with adverbs and formulaic terms are still regarded as variations, because adverbs and formulaic terms are often used to enhance or emphasize emotions or attitudes, but not to change the meaning of the whole utterance. *That's really true* and *I think that's certainly true* are cases in point, where *really* and *certainly* are adverbs, and *I think* is the formulaic term. They are used to emphasize the attitude of the speaker. Formulaic terms refer to expressions such as "*I think*" and "*I believe*", which display in the form of "I + predicate", to express the speaker's subjectivity in spoken discourse

---

[2] In the Switchboard Dialogue Act Corpus, restarts and non-sentence elements also have been marked within each utterance, such as filler ({F…}), discourse marker ({D…}) and coordinating conjunction ({C…}) (Meteer et al., 1995). In Example 2, "*You know*" is coded as discourse marker and "*Uh*" as filler.

(Baumgarten and House, 2010). Also, they have been recognized as one type of "engagement", dealing with "sourcing attitudes and the play of voices around opinions in discourse" in the appraisal framework (Martin and White, 2005: 35).

- Thirdly, the negative form and interrogative form, e.g. *that's not true*, *is that true?* are excluded, since their meanings and primary functions are apparently distinct from those of *that's true*.
- Fourthly, cases subsequently followed by that-clauses or prepositional phrases are excluded from the current work either, for instance

Example 3
sv   B.115 utt2: {C So } I do think <u>it's right</u> that they're harder on themselves, # {D you know. } # /
                                    sw_0382_4785.utt

It's true, followed by a that-clause, is not used independently any more. Such cases are not concerned with at this moment.

- Finally, it is necessary to reconsider the independent token *right* since it is often used as acknowledging token in the literature (e.g. Gardner, 2004; 2007), different from *that's right*. As a consequence, *right* is not treated as a variant of *that's right* in this stage, which will be verified by the statistical information later.

Variations of *that's right*	Variations of *that's true*	Variations of *that's correct*
	True	*Correct*
Adverb + *right*	Adverb + *true*	Adverb + *correct*
That's + adverb + *right*	*That's* + adverb + *true*	
Formulaic term + *that's* + *right*	Formulaic term + *that's true*	Formulaic term + *that's correct*
Formulaic terms + *that's* + adverb + *right*	Formulaic term + *that's*+ adverb + *true*	
It's right	*It's true*	
	It's + adverb + *true*	
	Formulaic term + *it's true*	
	Formulaic term + *it's* + adverb + *true*	
	This is true	
	This is + adverb + *true*	

Table 1 Variations of *that's right/true/correct*

Thus, the final list can be identified as shown in Table 1, where similar patterns take one-to-one correspondence. Apparently, *that's true* has more different types of variations than the other two.

4   Descriptive Statistics

That's right, *that's true* and *that's correct* are in effect synonymous concerning the dictionary meaning, while in the corpus, they do vary regarding their frequency information.

	(1) That's right and variations	(2) That's true and variations	(3) That's correct and variations
Total	911	920	21

(1), (2) and (3) in the following will be used to stand for the three sets of utterances respectively.

Table 2 Statistical information of the three sets

It is obvious that the total occurrence of (1) and (2) are almost the same, both of which far exceed that of (3). Beyond this, a range of functions have been identified for each of them, of which "aa", "ba", "s", "na" and "b"[3] are the most significant ones, all together accounting for over 98% in each set. Table 3 sets out these functions and their relative frequencies in performing each of them.

	aa	ba	s	na	b	Total
(1)	682 75%	139 15%	30 3%	26 3%	20 2%	897 98%
(2)	659 72%	148 16%	96 10%	2 0.2%	4 0.4%	909 99%
(3)	13 62%	5 24%	1 5%	1 5%	1 5%	21 100%

aa = accept; ba = assessment/appreciation; s = statement; na = affirmative answer; b = acknowledgement/backchannel

Table 3 Top five functions of three sets

[3] In the coding scheme SWBD-DAMSL, there are very specific definitions for each of them. *Accept* (aa), one subtype of *agreement*, indicates the speaker explicitly accepts a proposal, or makes agreements with previous opinions (Jurafsky et al., 1997: 37). *Assessment/appreciation* (ba) is defined as "a backchannel/continuer which functions to express slightly more emotional involvement and support than just 'uh-huh'" (Jurafsky et al., 1997: 48). *Statement* (s) divides into *"descriptive/narrative/personal"* statements (sd) and *"other-directed opinion statements"* (sv), both with the primary purpose of making claims about the world (including answers to questions) (Allen and Core, 1997: 10). *Affirmative answer* (na) is one subclass of *answers*, which indicates affirmative answers that are not "yes" or a variant (Jurafsky et al., 1997: 50). *Acknowledgement* (b) is usually "referred to in the CA literature as a 'continuer'" (Jurafsky et al., 1997: 42).

A glance at the table establishes that these top five functions together account for a large proportion among a series of functions performed by each particular set. In particular, *accept* overwhelmingly occurs in all the three sets, followed by *assessment/appreciation*. However, set (3) displays some slight distinction from (1) and (2) in the way that its proportion of *assessment/appreciation* is around 10% higher than that of sets (1) (2), but approximately 10% lower in *accept*. In the description to follow, the major concern is to seek similarities and distinctions within each set.

4.1   *That's right* and its variations

That's right and its variations frequently occur in daily speech, which can be seen in Table 4.

Types	Freq.	Percentage
That's right	852	93.5%
That's + adverb + *right*	26	2.9%
Formulaic term + *that's* + *right*	20	2.2%
Formulaic term + *that's* + adverb + *right*	5	0.6%
It's right	4	0.4%
Adverb + *right*	4	0.4%
Total	911	100%

Table 4 Statistical information of set (1)

It is perceptible that the simple token *that's right* overwhelmingly occurs compared to a range of variations, which may be indicative of the significance of economy in casual talk. By contrast, formulaic terms and adverbs are not so often attached with *that's/it's right*, accounting for less than 3% (2.2%+0.6%) and 4% (2.9%+0.6%+0.4%) respectively, which implies that such additional emphasis of stance and attitudes is not common in daily conversation. Noticeably, *it's right* appears 4 times, and *this is right* never occurs in the corpus. Hence *that*, *it* and *this* are similar lexical items but they have their own particular preference in some circumstance: when prefacing "be + right", *that* is more often used than *it* and *this*.

Regarding a variety of functions they serve, *that's right* and its variation totally perform twelve different functions in the corpus, but the top five are extremely significant which can be seen in Table 5, together constituting over 60% in each row. Strikingly, *that's right* does exhibit some slight distinction from its variations in that *that's right* can respond to a prior question and acknowledge to

what has been uttered, while its variations cannot do so.

Types	aa	ba	s	na	b	Total
That's right	642 76%	134 16%	19 2%	26 3%	20 2%	841 99%
That's + adverb + *right*	21 78%	3 15%	1 4%	0	0	25 96%
Formulaic term + *that's* + *right*	13 65%	1 5%	6 30%	0	0	20 100%
Formulaic term + *that's* + adverb + *right*	2 40%	0	1 20%	0	0	3 60%
It's right	0	1 25%	3 75%	0	0	4 100%
Adverb + *right*	4 100%	0	0	0	0	4 100%

aa = accept; ba = assessment/appreciation; s = statement; na = affirmative answer; b = acknowledgement/backchannel

Table 5 Top five functions performed by set (1)

Moreover, when formulaic terms are attached previously, the whole utterance has greater likelihood to function as *statement*. It is noted that the top three functions of *that's right* are exactly those functions analyzed and discussed in the literature, that is, agreement, assessments and affirmative answers. But with the empirical evidence, it can be further observed that agreement is much more remarkable than the other two. In addition, *it's right* is a special token in the table in that it clearly prefers *statement* to *accept*, which should not have been counted as a variant of *that's right*. Yet, considered the limited occurrence (4 times), it is not pervasive enough to determine what kind of functions it exactly serves, so it remains in this set. In the future, a larger spoken corpus will be in demand for examining such tokens.

4.2 *That's true* and its variations

Likewise, Table 6 exhibits basic frequency information of set (2). Different from set (1), *that's true* has much more variations than *that's right* in terms of types and tokens, which is illustrated by the statistics that variations of *that's true* make up 29% of set (2) while those of *that's right* just accounts for 6.5% of set (1). It needs to be noted that the symbol "*" in Table 6 means that the adverb in *that's* + adverb + *true* is able to move freely, not restricted to the middle position, such as "*probably that's true*", or "*that's true also*". This, however, has not been perceived for *that's right*.

Types	Freq.	Percentage
That's true	653	71.0%
*That's + adverb + *true*	93	10.1%
True	59	6.4%
Adverb + *true*	13	1.4%
Formulaic term + *that's true*	36	3.9%
Formulaic term + *that's* + adverb + *true*	11	1.2%
It's true	25	2.7%
It's + adverb + *true*	8	0.9%
Formulaic term + *it's true*	2	0.2%
Formulaic term + *it's* + adverb + *true*	1	0.1%
This is true	15	1.6%
This is + adverb + *true*	4	0.4%
Total	920	100.0%

Table 6 Statistical information of set (2)

Yet still, set (2) is consistent with set (1) in two respects. On the one hand, *that's true* occurs more frequently than *it's true* and *this is true*, which is correspondingly close to set (1). On the other hand, formulaic terms and adverbs do not show high frequency in set (2) either. *That's true*, *it's true* and *this is true* are far more frequently used than those embedded with formulaic terms or adverbs.

Types	aa	ba	s	na	b	Total
That's true	487 74%	105 16%	55 8%	0	2 0.3%	649 99%
*That's + adverb + *true*	62 67%	15 16%	8 9%	2 2%	1 1%	88 96%
True	38 64%	18 31%	0	0	1 2%	57 97%
Adverb + *true*	10 77%	3 23%	0	0	0	13 100%
Formulaic term + *that's true*	22 61%	1 3%	13 35%	0	0	36 100%
Formulaic term + *that's* + adverb + *true*	6 55%	1 9%	3 27%	0	0	10 91%
It's true	11 44%	3 12%	11 44%	0	0	25 100%
It's + adverb + *true*	5 62.5%	0	3 37.5%	0	0	8 100%
Formulaic term + *it's true*	0	0	2 100%	0	0	2 100%
Formulaic term + *it's* + adverb + *true*	0	0	1 100%	0	0	1 100%
This is true	13 87%	2 13%	0	0	0	15 100%
This is + adverb + *true*	4 100%	0	0	0	0	4 100%

aa = accept; ba = assessment/appreciation; s = statement; na = affirmative answer; b = acknowledgement/backchannel

Table 7 Top five functions performed by set (2)

Concerning a range of functions they perform, *that's true* and its variations totally have nine different functions in the corpus, among which the top five are displayed in Table 7. Overall, the distribution here shares a large number of similarities with that of set (1) in Table 5. In particular, *accept*, *assessment/appreciation* and *statement* are considerably significant, while *affirmative answer* and *acknowledgement* are comparatively less crucial, only occurring in *that's true*, *that's* + adverb + *true* and *true*. When *that's true* is attached with formulaic terms, the likelihood to function as *accept* declines accompanying with greater proportion in *statement*. The exceptional token is *it's true*, which itself prefers both *accept* and *statement*. In this sense, *it's true* is distinguished from *that's true* which overwhelmingly deals with *accept*. By contrast, *this is true* is relatively consistent with *that's true* in primary functions they serve. Thus, in the pattern "THAT/IT/THIS + BE + TRUE", *that*, *it* and *this* indicate their particular preference as well.

4.3   *That's correct* and its variations

That's correct and its variations are used infrequently, with a total occurrence of 21 in the whole corpus. As a consequence, there are far less variations in this set. Table 8 shows the basic frequency information, and Table 9 exhibits all functions performed by these tokens.

Types	Freq.	Percentage
That's correct	13	61.9%
Formulaic term + *that's correct*	3	14.3%
Correct	4	19.0%
Adverb + *correct*	1	4.8%
Total	21	100.0%

Table 8 Statistical information of set (3)

Types	aa	ba	s	na	b	Total
That's correct	8 62%	4 31%	0	1 8%	0	13 100%
Formulaic term + *that's correct*	2 67%	0	1 33%	0	0	3 100%
Correct	2 50%	1 25%	0	0	1 25%	4 100%
Adverb + *correct*	1 100%	0	0	0	0	1 100%

aa = accept; ba = assessment/appreciation; s = statement; na = affirmative answer; b = acknowledgement/backchannel

Table 9 All functions performed by set (3)

As can be seen in Table 8, *that's correct* occurs more frequently than its variations, accounting for 62% in set (3), which is lower than that of *that's right* (93%) and *that's true* (71%). Moreover, formulaic terms and adverbs are not so frequent, either, which suggests bare tokens such as *that's correct* and *correct* are preferred by native speakers. Considered a range of functions performed Table 9, *accept* and *assessment/appreciation* are remarkable compared to *statement*, *affirmative answer* and *acknowledgement* with one occurrence for each.

To summarize, an overview of utterances in the three sets has presented with empirical evidence. Generally, they share quite a lot of similarities in terms of primary functions they serve. In addition, two points need to be further elaborated. One is that, *right* is assumed to be used in a way different from *that's right* in conversation, which is further confirmed by the evidence that 73% of *right* serve *acknowledgement* while *that's right* prefers *accept* with 76% of its total occurrence in the corpus. This can be observed in Table 10, where their top five functions have been listed respectively. Also, 16% of *that's right* can be used as *assessment/appreciation*, whereas the single *right* only occurs 11 times (0.2%) as *assessment/appreciation*. Hence, in general, *right* and *that's right* are two different cases in interactive speech.

	b	aa	na	%	fc	Total
Right	3685 73%	1154 23%	127 3%	26 0.5%	17 0.3%	5009 99%
	aa	ba	na	b	s	Total
That's right	642 76%	134 16%	26 3%	20 2%	19 2%	841 99%

aa = accept; ba = assessment/appreciation; s = statement; na = affirmative answer; b = acknowledgement/backchannel; % = abandoned utterances; fc = conversational closing

Table 10 *right* vs. *that's right*

The second point is that, *that*, *it* and *this* have their particular preference to the pattern "THAT/IT/THIS+BE+RIGHT/TRUE/CORRECT", which can be summarized as follows.

That > Ø > it > this

It means that the ones on the left side take priority over those on the right: *that* more likely occurs than *this*, and the symbol Ø signals no pronoun occurs. This is highly consistent with Tao's finding (2003: 202) "*that* is more likely to be used as a turn initiator than *this*".

5   A Comparative Study

According to the previous statistical analysis, it is noted that *that's right*, *that's true* and *that's correct* account for quite a large proportion in each particular set. The previous observation has also shown that the total occurrence of *that's correct* is much fewer than the other two, and therefore, a comparative study will concentrate on *that's right* and *that's true*, and examine the condition where they are mutually exchangeable with each other and where they are distinct from each other. Figures 1 and 2 respectively fill out their primary functions and their preceding contexts[4]. By Figure 1, apparently *that's right* and *that's true* both exhibit considerable preference to *accept* and *assessment/appreciation* which together make up over 90% for both cases. It is meant that over 90% of their tokens perform the two same functions.
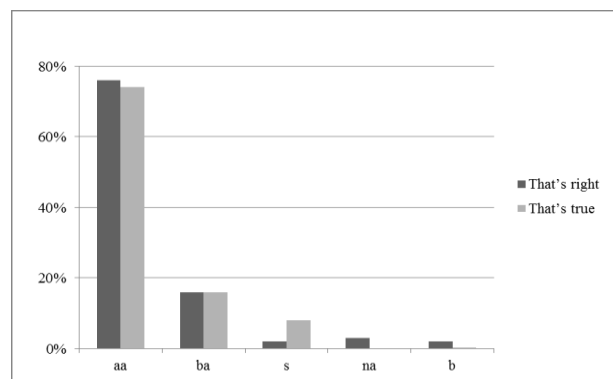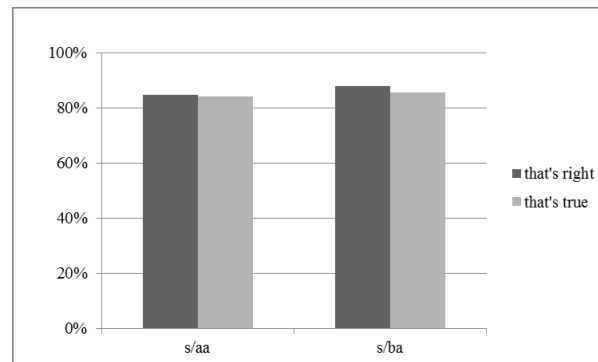

aa = accept; ba = assessment/appreciation; s = statement; na = affirmative answer; b = acknowledgement/backchannel

Figure 1 Primary functions of *that's right* and *that's true*

However, some slight difference between them can be perceived as well. *That's right* is used to perform all these five functions, while *that's true* cover four of them and cannot be not used to answer a question. At the same time, *that's true* shows far greater likelihood to serve *statement* compared to *that's right*. By contrast, *that's right* is almost ten times more likely than *that's true* to function as *acknowledgement*.

In order to see whether their previous contexts could offer useful cues to differentiate the occurrence of *that's right* and *that's true*, a specific view is taken into the previous contexts when they act as

accept and assessment/appreciation, because the two functions together make up a large proportion of the total occurrence. Figure 2 depicts the salient previous context when they act as the two functions.

aa = accept; ba = assessment/appreciation; s = statement

Figure 2   Previous contexts of *aa* and *ba*

It is clear that *statement* is the most overwhelming previous function, accounting for over 80% previous contexts of *that's right/true* when they act as *accept* and *assessment/appreciation*. It seems that the previous contexts offer little cues to differentiate them, since both are so often preceded by *statement*. According to Figures 1 and 2, it is possible that almost 75% of *that's right/true* are mutually exchangeable since over 90% of their occurrence contributes to *accept* and *assessment/appreciation*, in which over 80% of the previous contexts are *statement*. This can be further validated by the chi-square test, which aims to test if *that's right* and *that's true* have no difference in the distribution of different functions. Table 11 shows the frequency distribution of *that's right/true* in *accept*, *assessment/appreciation* and other functions. Table 12 exhibits the result of the test.

			Functions			Total
			aa	ba	others	
To-kens	That's right	Count	642	134	76	852
		Expected count	639.1	135.3	77.6	852.0
	That's true	Count	487	105	61	653
		Expected count	489.9	103.7	59.4	653.0
Total		Count	1129	239	137	1505
		Expected count	1129.0	239.0	137.0	1505.0

aa = accept; ba = assessment/appreciation

Table 11   Tokens*functions crosstabulation

---

[4] The previous contexts are restricted to immediately previous utterances uttered by others.

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.130[a]	2	.937
Likelihood Ratio	.130	2	.937
N of Valid Cases	1505		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 59.44.

Table 12   Chi-Square Tests

In Table 12, the value of pearson chi-square is 0.130, and the p-value is 0.937 which is larger than 0.05. It manifests that the difference between *that's right* and *that's true* is not significant in the distribution of primary functions according to the frequency information observed in the corpus.

In summary, the statistical analysis above demonstrates that *that's right* and *that's true* are used almost the same in interactive speech, in which nearly 75% of their total occurrence are interchangeable. This is further confirmed by the significant test which explicitly shows no significance in the distribution of primary functions, and their previous contexts supply little cues for the distinction. In some cases, however, they have their own preference and differ from each other. For instance, *that's true* has never been found to answer a previous question in the corpus, while 3% of *that's right* can perform this function. Moreover, *that's true* shows much greater likelihood to serve *statement* whereas *that's right* is almost ten times more likely than *that's true* to be *acknowledgement*. Specifically, when the preceding utterance is a statement or a question, the current utterance is more likely to serve *statement* if it is realized by *that's true*; it has greater possibility to be *acknowledgement* or *an affirmative answer* if it is realized by *that's right*. This kind of preference is expected to facilitate DA tagging.

6   Conclusions

This paper presented a quantitative investigation of three short utterances (i.e. *that's right*, *that's true*, *that's correct*) and their variations in the Switchboard Dialogue Act Corpus. Particularly, it offered an overview to account for how they are used in daily conversation with empirical evidence. By the current investigation, it has been observed that *that's right/true* and their variations much more frequently occur than *that's correct* and its variation. In terms of primary functions served in interactive speech, they consistently exhibit great

preference to *accept, assessment/appreciation, statement, affirmative answer* and *acknowledgement*, among which, *accept* and *assessment/appreciation* together account for quite a large proportion. Regarding their variations, *that*, *it* and *this* are similar lexical items but they indicate their particular preference to this pattern "THAT/IT/THIS+BE+RIGHT/TRUE/CORRECT" . Moreover, formulaic terms and adverbs are not so frequently embedded. When formulaic terms are attached, the whole utterances have greater likelihood to be *statement*.

Also, we have specified some crucial issues for *that's right* and *that's true*, which are clearly useful to the detection of DAs. It has been discovered that almost 75% of *that's right* and *that's true* are mutually exchangeable, which has been verified by the chi-square that their difference is not significant in the distribution of primary functions. Moreover, the previous contexts offer little cues to differentiate *that's right* and *that's true*. In this sense, they are two short utterances with similar meanings and uses. But in some cases, they display their particular preference: *that's right* has fewer variations compared to *that's true*, and covers a wide range of functions in the corpus; *that's true* has never been found to answer a previous question in the corpus, while 3% of *that's right* can do that. Moreover, *that's true* shows much greater likelihood to serve *statement* whereas *that's right* is more likely to be *acknowledgement*. Such kind of empirical analysis will provide the insights and bases for automatic DA tagging. In addition, we believe that it also tells second language learners how to use these three shore utterances under specific contexts.

Acknowledgements

References

Allen, J. & M. Core. 1997. DAMSL: Dialogue Act Markup in Several Layers (Draft 2.1). Technical Re-

port, Multiparty Discourse Group, Discourse Resource Initiative, September/October 1997.

Baumgarten, N., & House, J. 2010. I think and I don't know in English as lingua franca and native English discourse. Journal of Pragmatics, 42(5), 1184-1200.

Bunt, H., Alexandersson, J., Carletta, J., Choe, J.-W., Fang, A.C., Hasida, K., Lee, K., Petukhova, V., Popescu-Belis, A., Romary, L., Soria, C. and Traum, D. 2010. Towards an ISO standard for dialogue act annotation. In Proceedings of the Seventh International Conference on Language Resources and Evaluation. Valletta, MALTA, 17-23 May 2010.

Fang, A., H. Bunt, J. Cao, and X. Liu. 2011. Relating the semantics of dialogue acts to linguistic properties: A machine learning perspective through lexical cues. In Proceedings 5th IEEE International Conference on Semantic Computing, Stanford University, Palo Alto.

Gardner, R. 2001. When Listeners Talk: Response Tokens and Listener Stance. John Benjamins Publishing

Gardner, R. 2004. Acknowledging strong ties between utterances in talk: Connections through right as a response token. In Proceedings of the 2004 Conference of the Australian Linguistic Society, pages 1–12.

Gardner, R. 2007. The Right connections: Acknowledging epistemic progression in talk. Language in Society, 36(03), 319-341.

ISO Standard 24617-2. 2012. Language resource management–Semantic annotation frame-work (SemAF), Part 2: Dialogue acts. ISO, Geneva, 2012.

Jurafsky, D., Shriberg, E. and Biasca, D. 1997. Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual, Draft 13. University of Colorado, Boulder Institute of Cognitive Science Technical Report 97-02.

M. Meteer and A. Taylor. 1995. Dysfluency annotation stylebook for the Switchboard Corpus. available online at ftp://ftp.cis.upenn.edu/pub/treebank/swbd/doc/DFL-book.ps.

Martin, J. R., & White, P. R. 2005. The language of evaluation. Basingstoke and New York: Palgrave Macmillan.

Mayor, M. (Ed.). 2009. Longman dictionary of contemporary English. Pearson Education India.

McCarthy, M. 2003. Talking back: "Small" interactional response tokens in everyday conversation. Research on Language and Social Interaction, 36(1), 33-63.

Stenström, A. B. 1987. Carry-on signals in English conversation. Corpus linguistics and beyond, 87-119.

Tao, H. 2003. Turn Initiators in Spoken English: A Corpus-Based Approach to Interaction and Grammar. Corpus Analysis: Language Structure and Language Use, (46), 187.

Tsui, A. B. 1994. English conversation. Oxford Univ Pr.