

Equative and Predicational Copulas in Thai

NANCY HEDBERG^a, DAVID POTTER^b
Simon Fraser University^a; Northwestern University^b

Introduction

In this paper we report on an experimental study of the copula system in Thai. We argue that the Thai copula system is crucial to the debate over whether specificational sentences should be given an inverse predicational analysis or an equative analysis. We conclude that Thai data supports the equative approach. The paper is organized as follows. We first introduce the inverse and equative analyses, and provide a little background on Thai. We then discuss our sentence acceptability task survey, and introduce the idea that the test items in the survey may have been subject to type-shifting. After that, we describe some interviews we did with Thai speakers aimed at fixing the semantic types of the test items. Finally, we discuss the results of the interviews and conclude.

1 The Inverse vs. Equative Controversy

Higgins 1973 presented a classification of copular sentences in English into four types of interpretation, as shown in (1). The examples here are taken from Mikelsen 2005.

- (1) a. Predicational: *Susan is a doctor*
- b. Specificational: *The winner is Susan*
- c. Identity: *She is Susan*
- d. Identificational: *That woman is Susan*

The controversy has to do with which of Higgins' categories should be grouped together into different senses of the copula. The proper analysis of specificational sentences is at the heart of the controversy.

The Inverse Analysis has a number of recent adherents, including Williams 1983, and Partee 1986 for English, Moro 1997 for Italian, Frank 2002 for English,

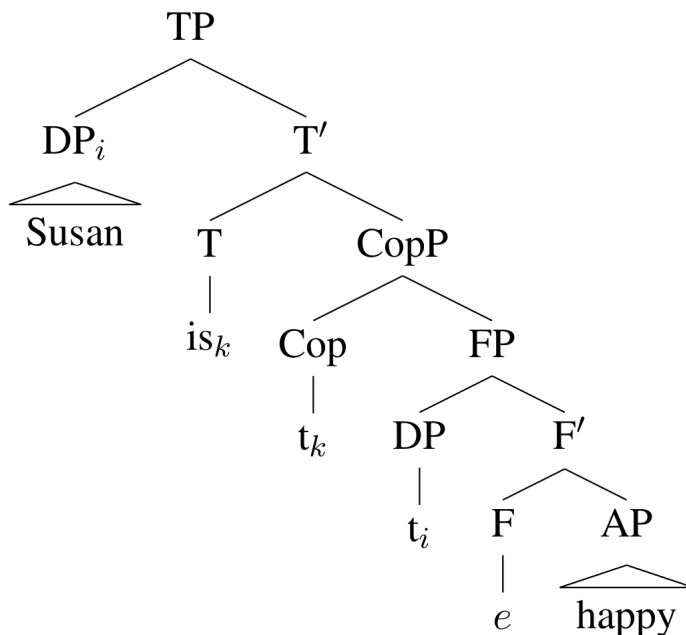
Adger and Ramchand 2003 for Scots Gaelic, Mikkelsen 2005 for English and Danish, and den Dikken 2006 for English and Dutch.

This approach takes specificational and predicational sentences to derive from the same base, typically taken to be a small clause. Predicational sentences result from raising the subject of the small clause, an *e* type argument, to subject position of the sentence, as in (2). In contrast, specificational sentences result from raising the predicate of the small clause, an $\langle e, t \rangle$ type argument, as in (3). In this analysis then, reverse specificational sentences are just predicational sentences with subject focus, as in (4).

- (2) A: What qualities does Susan have?
 B: Susan is the winner of the student prize.
- (3) A: Who is the winner of the student prize?
 B: The winner of the student prize is Susan.
- (4) A: Who is the winner of the student prize?
 B: SUSAN is the winner of the student prize.

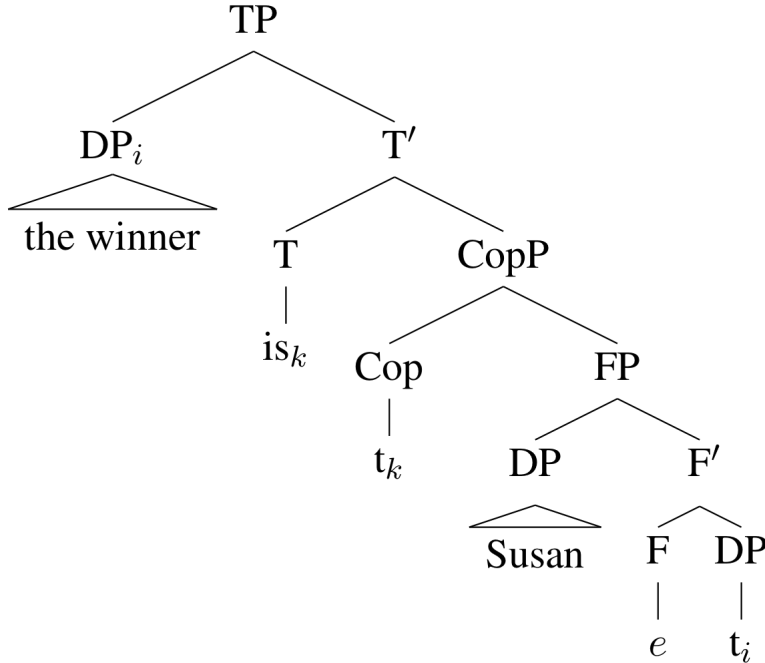
The trees in (5) and (6) show a sketch of how this analysis of predicational and inverse sentences works, based on the syntactic analyses of Frank 2002. FP here is a small clause headed by some functional category.

(5) Predicational Tree



Predicational sentences result from raising the subject of the small clause; and specificational sentences, called ‘inverse’ sentences, result from raising the predicate of the small clause. Reverse specificational sentences like (4b) are treated as instances of predicational sentences, with a definite DP in predicate position of the small clause where the AP is in (5).

(6) Specificational tree on inverse approach



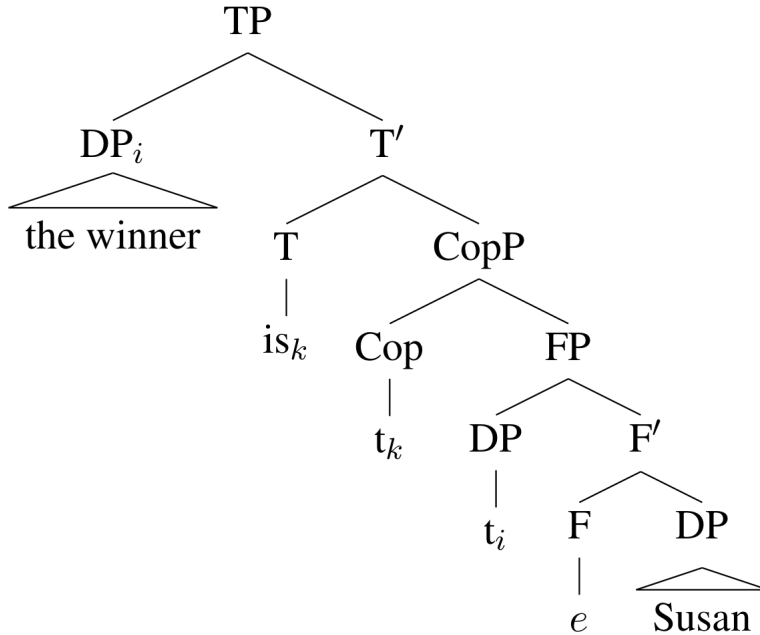
Thus, on the inverse approach, predicational, specificational and reverse specificational sentences receive identical semantic interpretations, and all result from the same syntactic base containing a vacuous copula that takes a small clause complement. Either the subject or the predicate of the small clause raises to the external subject position (specifier of TP).

A second copular analysis is the equative approach, taken by Zaring 1996 for Welsh, Heycock and Kroch 1999 for English, and Han and Hedberg 2008 for clefts in English. On this approach specificational and predicational sentences are derived from a distinct bases; both specificational and identity sentences are involve an equative copula. Han and Hedberg assume that the descriptive argument in specificational sentences is a generalized quantifier (Barwise and Cooper 1981), i.e. a definite description interpreted quantificationally, and that their semantic interpretation involves the equals sign. These sentences, then, receive the Russellian interpretation shown in (7), where the subject is of type $\langle\langle e, t \rangle, t \rangle$ and the complement is an e type argument.

$$(7) \quad \exists x [\text{winner}(x) \wedge \forall y[\text{winner}(y) \rightarrow y=x] \wedge x = \text{Susan}]$$

A sketch of a syntactic analysis for the specificational sentence on the equative approach, based on Han and Hedberg 2008, is shown in (8).

(8) Specificational tree on equative approach



On the equative approach that we propose, reverse specificational sentences result from the generalized quantifier $\langle\langle e,t \rangle, t \rangle$ type argument occupying the complement position while the e type argument occupies the subject position. Thus, specificational sentences are viewed as reversible, like many identity and identificational sentences. Identity sentences like that in (1c) and identificational sentences like that in (1d) make use of the same equative copula as specificational sentences ($\lambda y \lambda x [x=y]$) but both subject and complement are of type e .

Predicational sentences on the equative approach receive the same syntactic analysis as on the inverse approach, as illustrated in (5), and they are interpreted as containing an initial e type (or $\langle\langle e,t \rangle, t \rangle$ type) argument and then an $\langle e,t \rangle$ argument.

Thus, under the equative analysis, predicational, specificational, and reverse specificational sentences involve a small clause, the FP, and, in contrast with the inverse analysis, it is always the subject of the small clause that raises. It is therefore the semantics of these sentences that distinguishes them: specificational semantics result from a non-vacuous, equative copula and a quantificationally-interpreted DP; while predicational semantics follow from a vacuous copula and a predicatively-interpreted AP or DP.

2 Background on Thai

We now provide some background on Thai. Kuno and Wongkhomthong 1981 discuss the interpretation of sentences with two different copulas in Thai, *bpen* and *kheuu*, ‘be,’ as illustrated in (9). They label sentences containing the copula *kheuu* as ‘identificational’ sentences.

- (9) a. yâpùn **pen/*khi:** pràthê:d ùdsa:hàkam.
 Japan is country industry
 ‘Japan is an industrial country.’
- b. khon thî: chán rág ***pen/khi:** cɔ:n
 person that I love is John
 ‘The person that I love is John.’

About (9b), they say “[9b] is a sentence that identifies the person that the speaker likes best with John. It cannot be interpreted as a sentence that presents as one of the characteristics of the person that the speaker likes best the fact that he is John. Hence, the sentence is exclusively identificational....” *Bpen*, on the other hand appears in ‘characterizational’ sentences like (9a).

Either copula can appear in (10). They say, “If the speaker’s intention is to present one of John’s characteristics, *pen* is used. On the other hand, if the speaker’s intention is to equate John and the person he (= the speaker) likes best, then *khi:* is used.”

- (10) cɔ:n **pen/khi:** khon thî: chán rág
 John is person that I love
 ‘John is the person that I love.’

Kuno and Wongkhomthong explicitly relate *bpen* to Higgins’ predicational sentences and *kheuu* to his specificational and identificational sentences. We agree with this analysis, and hypothesize that, consistent with the equative analysis, *bpen* appears in predicational sentences and *kheuu* appears in specificational, identificational and identity sentences. Kuno and Wongkhomthong’s discussion is descriptively very rich and convincing. We designed our study to confirm their analysis by presenting relevant data to large groups of participants, to examine some gaps in the data they discussed, and to bring the Thai data to bear more concretely on the current copula sentence controversy.

Specifically, we argue that Thai distinguishes the copula used in predicational and specificational sentences along the lines that the equative analysis predicts (e.g. the distinction between (9a) and (9b)). We also argue that reverse specificational sentences need to be distinguished from predicational sentences, along the lines of the difference between the two copulas exemplified in (10). Crucially,

kheuu sentences with the two possible word orders receive the same interpretation, while *bpen* sentences with only one word order receive a different interpretation.

3 The Experimental Survey

We now introduce our experimental survey, which utilized a sentence acceptability judgment task. The stimuli presented both copulas, each in one of four sentence contexts, shown in (11). We included a specificational context, an identity context, and two predicational contexts. Thus, there were eight conditions all together.

In order to try to ensure a specificational instead of an identificational interpretation of the sentences, we used a description with no determiner but with a superlative modifier. In this way, we attempted to invoke a quantificational definite description reading associated with uniqueness (type $\langle\langle e,t\rangle,t\rangle$) instead of a demonstrative referential reading associated with familiarity (type e). Our stimuli here consisted of occupational nouns with superlative modifiers. We used adjectival free relatives with a classifier instead of simple adjectives because simple adjectival sentences in Thai don't contain either copula.

- (11) a. **Specificational:** Definite Description + *bpen/kheuu* + Proper name

Nak a daaeng	khohn thee	chan	tai ruup	maak thee soot
Actor	person that	I	photograph	<i>superlative</i>
<i>bpen/kheuu</i>	khoon		Angela	
is	<i>politeness-marker</i>	Angela		

‘The actor that I photograph most is Angela.’

- b. **Identity:** Pronoun + *bpen/kheuu* + Proper name

Khao	<i>bpen/kheuu</i>	khoon	Angela
He/she	is	<i>politeness-marker</i>	Angela

‘She is Angela.’

- c. **Predicational (1):** Definite Description + *bpen/kheuu* + Adjectival Free Relative

Nak sa daaeng	khohn thee	chan	tai ruup	maak thee soot
Actor	personthat	I	photograph	<i>superlative</i>
<i>bpen/kheuu</i>	khohn	suay		
is	person	pretty		

‘The actor that I photograph most is a pretty person.’

Equative and Predicational Copulas in Thai

d. **Predicational (2):** Pronoun + *bpen/kheuu* + Adjectival Free Relative

Khao *bpen/kheuu* khohn suay
He/she is person pretty
'She is a pretty person.'

Out of a total of 120 proper names, 120 descriptions and 120 adjectival modifiers, a pool of 960 stimulus sentences was created, 120 in each of the eight conditions. Eight test versions were created. Each test contained 120 test stimuli. Subject and object tokens were counterbalanced across test versions, and each of the eight conditions were presented 15 times in each test. Each test also contained 120 fillers: 60 grammatical sentences and 60 ungrammatical sentences. Fillers were independently rated grammatical or ungrammatical by three native speakers. Also, while the grammatical sentences were perfectly grammatical, the ungrammatical sentences ranged from 'word soup' to mildly ungrammatical.

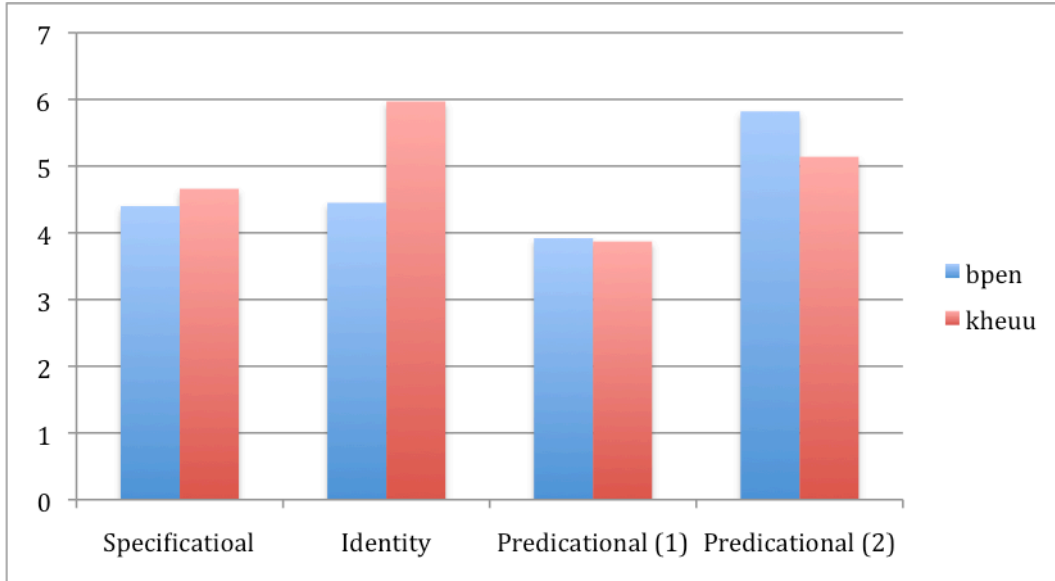
The participants were 49 native Thai speakers, who were students recruited from Chukalongkorn and Suan Dusit Universities in Bangkok, Thailand. We report on the results of 40 participants since nine participants were excluded as their mean rating of the ungrammatical sentences was above 4.

The procedure was as follows. The task was a written acceptability judgment task, with acceptability characterized as how 'natural' or 'normal' each sentence sounds to the participant. There was a seven point rating scale, ranging from 1 'completely unacceptable' to 7 'completely acceptable'. Task materials and instructions were given in Thai. Participants were advised to judge based on their first impression and to not second-guess their judgments.

The experimental hypotheses were the following. (1) In specificational contexts, *kheuu* should be more acceptable than *bpen*. (2) In identity contexts, *kheuu* should be more acceptable than *bpen*. (3) In predicational contexts, *bpen* should be more acceptable than *kheuu*.

A summary of the results are presented in (12). For each condition, a one-way analysis of variance (ANOVA) was performed with copula as a repeated measure. In specificational contexts, *kheuu* did turn out more acceptable than *bpen*. This difference appears very close, but was statistically significant (*bpen* (4.40), *kheuu* (4.66), [F(1,39) = 4.565, $p < .05$]). In identity sentences, *kheuu* was again more acceptable than *bpen* and the difference was statistically significant (*bpen* (4.45), *kheuu* (5.97), [F(1,39) = 45.12, $p < .001$]). In the predicational (2) context (the one with a pronoun subject), *bpen* was more acceptable than *kheuu* and the difference was statistically significant (*bpen* (5.82), *kheuu* (5.14), [F(1,39) = 26.98, $p < .001$]). There was no significant difference between the copulas in the predicational (1) context.

(12) Results of Experimental Survey



While our hypotheses were supported statistically, we were surprised at how close the results were, and we were puzzled at why sentences with the non-preferred copula were judged acceptable to such a great degree. We have thought about two different types of explanation for the closeness of the results: problems with the acceptability of our particular stimuli, and also the possibility of type shifting in the interpretation of our sentences.

With regard to the first possibility, we note that the specificational and predicational (1) sentences were somewhat awkward in that they convey more than one proposition. We don't know how the participants judged the sentences, and the availability of a simpler and equivalently interpreted alternative discourse sequence may have reduced the rating of the test items. For example for the specificational sentence in (11a) it would be simpler to split the sentence into two: "Angela is an actor. 'I photograph her most'". Likewise for the predicational sentence in (11c), it might be easier to understand if the sentence were split into two, i.e. "I photograph one actor most. She is a pretty person".

Note also that these two types of sentence were longer than the identity and predicational (2) sentences. Thus, the absolute ratings of these conditions are not comparable, and the length difference may have obscured the judgments of the participants.

The results may also have been influenced by a number of non-truth-conditional factors. For example, more than one participant noted that *kheuu* is strongly preferred in formal writing, and that it signifies a greater confidence in the truth of the statement.

Importantly however, we also decided that the sentences were possibly susceptible to type shifting of the arguments and that this might have affected the results. The differences were mostly less than 1 point on a 7-point scale. Could participants have been construing the sentences in different ways? After all, copular sentences are well known for supporting different construals, as discussed for example in Higgins 1973.

Perhaps the particular sentence elements can be shifted into different semantic types (Partee 1987). In particular, the adjectival free relatives consist of a classifier and an adjective. We were assuming that they were interpreted as predicates of type $\langle e, t \rangle$ but it is possible that they can be interpreted as referential (type e) or as generalized quantifiers (type $\langle \langle e, t \rangle, t \rangle$)—that is as ‘the pretty person’. Also the definite descriptions, which we had assumed were interpreted as referential or as generalized quantifiers could possibly have been interpreted as predicates (type $\langle e, t \rangle$), especially if occurring in complement position.

While an explanation for the closeness of the acceptability judgments in terms of type-shifting is possible for predicational (1) and predicational (2) contexts, we didn’t expect it to have an effect on the identity context or the specificational context since in those cases the proper name complement would seem to be an unshiftable argument of type e . In order to test the hypothesis concerning type shifting in the predicational contexts, we conducted some interviews with native speakers. We also wanted to extend the domain of our investigation to include reverse specificational contexts in which definite descriptions occur as copular complements. We discuss these interviews in the next section.

4 The Interviews

Four native Thai speaking professors and one businessman were interviewed in English. The notion of grammaticality was explained, with examples given in Thai. Then the crucial Thai sentence types were presented for grammaticality judgment, first out of context and then in two English contexts each. The contexts were designed to fix the semantic types of the test items.

The first question concerned the Predicational (1) sentence: *Actor that I like best kheuu/bpen khohn tall*, with contexts shown in (13).

- (13) **Context A:** “Suppose that you are describing a number of qualities about your favorite actor. She is sweet, brunette, Canadian and witty. Then you say that the actor you like best is tall.”

Context B: “Out of a room full of actors, you are asked to identify the actor that you like the most. You identify the tallest person in the room as your favorite actor.”

Out of context, 5/5 participants preferred *bpen*. We felt that Context A supported an $e + \langle e, t \rangle$ predicational interpretation, and indeed all 5 participants preferred *bpen*. However, in Context B, we felt that type-shifting would apply, turning the response into a generalized quantifier + referential—i.e. specificational—sentence: *Actor that I like best is the tall person*, and indeed 5/5 participants preferred *kheuu* here. Thus, type-shifting could explain why there was no significant preference for either copula in this sentence context in the survey.

The second question concerned the Predicational (2) sentence: *He kheuu/bpen khohn tall*, with contexts shown in (14).

- (14) **Context A:** “You witness a crime. The police officer asks you to describe the culprit. You list a number of traits, including that he is tall.”

Context B: “You witness a crime. Out of a police lineup, you are asked to identify the culprit. You identify the tallest person as him.”

Out of context *bpen* was preferred by 5/5 participants. Context A was indeed a predicational context: 5/5 participants preferred *bpen*, and we suggest that this supports an $e + \langle e, t \rangle$ interpretation of the sentence. However, we suggest that type shifting would occur in Context B, so that the sentence would be interpreted as ‘he/it is the tall person’, with an e or $\langle \langle e, t \rangle, t \rangle$ type subject and an e -type complement—i.e. an identity sentence. In support of this, note that 5/5 participants preferred *kheuu* here.

The third sentence type included in the interviews had not been included in the survey: a reverse specificational sentence—*Mr Sun kheuu/bpen actor that I like best*, with contexts shown in (15).

- (15) **Context A:** “Suppose that you are talking about the famous actor Mr. Sun. You say that he is short, fat and has a reputation for being short tempered. Nevertheless, you continue your description by saying that Mr. Sun is the actor that you like best.”

Context B: “I’m asking about different people in your life. I ask who your mother is and who your best friend is. Then I ask who the actor that you like best is.”

Out of context both copulas were accepted, with subjects unable to explain any meaning difference. We hypothesized that Context A would support a predicational interpretation ($e + \langle e, t \rangle$), and indeed, 5/5 participants preferred *bpen* here. However, we hypothesized that Context B would support a reverse specificational interpretation of the sentence and that *kheuu* would be preferred. This would be the assumption of the equative approach, which predicts that specificational

Equative and Predicational Copulas in Thai

sentences should be reversible, as well as that there should be at least a semantic difference between reverse specificational and predicational sentences, if not a morphosyntactic difference. Interestingly, it turned out that 4/5 participants did prefer *kheuu* although 1 preferred *bpen*. We conclude that 4/5 of the speakers were interpreting the sentence as $e + \langle\langle e, t \rangle, t \rangle$ rather than $e + \langle e, t \rangle$, thus supporting the need to distinguish between those two interpretations

The fourth, identity sentence type involved the sentence, *He kheuu/bpen Mr. Sun*, with contexts shown in (16).

(16) **Context A:** “You are introducing your friends to a group of people. You say, ‘She is Jill. He is John. He is Mr. Sun’.”

Context B: “Suppose that you find yourself face to face with King Rama (who can change his shape at will). At first he appears in his true form. Then suddenly, he changes to look like your best friend. Then he changes to look like Mr. Sun.”

Out of context, *kheuu* was preferred. We thought Context A supported a referential + referential interpretation, and *kheuu* was preferred by 4 out of 5 participants. The fifth speaker judged both copulas to be ungrammatical. Again we judged Context B to support an $e + e$ interpretation. Three participants preferred *kheuu* here. One participant preferred a ‘become’ type sentence with *bpen*, and one participant felt that both copulas were ungrammatical.

Finally, the specificational contexts shown in (17) involved the sentence, *Actor that I like best kheuu/bpen Mr. Sun*.

(17) **Context A:** “Out of a roomful of actors, you are asked to identify the actor that you like the most. You identify Mr. Sun.”

Context B: “Suppose that you are huge movie fan. “You have always had a favorite actor, but the identity of your favorite actor changes constantly. Last year, actor that I like best was Mr. Jones. Last month, actor that I like best was Mrs. Kim. Currently actor that I like best is Mr. Sun.”

Out of context, *kheuu* was preferred by 5/5 participants. Context A was intended to be a specificational context. *Kheuu* was preferred (5/5). We concluded that this context supported a generalized quantifier + referential interpretation. Context B was a variant of this. Again, *kheuu* was preferred (5/5), but *bpen* was allowed by two participants who said that it was marginal. We thus concluded that even this context supported a specificational interpretation.

However, one participant noted that, in context B, if the temporal modifier ‘now’ is included, then both copulas can be marginal, but without this modifier,

then it seems to be a statement of fact and *kheuu* is preferred. This seems to indicate that there is something else going on in the choice between *bpen* and *kheuu* besides semantic type, syntactic configuration and truth conditions. We intend to explore this possibility in future work.

To summarize our interpretations of the interviews, we conclude that our “definite descriptions” can be construed in terms of all three semantic types. Our “adjectival free relatives” can be construed as either referential or as predicates. *Bpen* is preferred in $e + \langle e,t \rangle$ contexts, and *kheuu* is preferred in $e + e, \langle \langle e,t \rangle, t \rangle + e$ and $e + \langle \langle e,t \rangle, t \rangle$ contexts.

The fact that *bpen* is preferred in context A and *kheuu* is preferred in Context B of the reverse specificational condition in (15) is strong evidence in favor of the equative analysis for both English and Thai because it is only the equative approach that predicts that reverse specificational sentences should be distinguished from predicational sentences in languages of the world—whether only semantically (as in English) or also morphosyntactically (as in Thai). If we consider the English interpretations of such sentences in relation to their Thai equivalents, they thus seem to be ambiguous in the way that the equative analysis predicts. The inverse analysis treats the English sentences as univocal. That is, it doesn’t allow for a reverse specificational interpretation distinct from a predicational interpretation since the complement argument in both cases can only be the predicate of a small clause with a vacuous copula selecting it. This approach thus cannot explain the difference between the Thai variants.

5 Conclusions

From the experimental survey, we conclude that the results support our experimental hypotheses. From the interviews we conclude that the small differences in the survey data can be explained to be the result of type shifting. Different participants construed the sentences in different ways, consistent with different ways of type shifting the sentence elements. A reverse specificational context was included in the interviews and favored *kheuu*, a finding which supports the equative analysis. Currently we are planning a second round of interviews to more systematically test speakers’ judgments in context, and in this round of interviews we will again include a reverse specificational sentence type.

Our theoretical conclusion is that the Thai data support the equative analysis. Two different copulas appear in predicational vs. equative sentence types. *Bpen* is preferred in predicational contexts. *Kheuu* is preferred in identity, specificational and reverse specificational contexts. More generally, we would like to argue that if one language (Thai) lexically and syntactically distinguishes predicational from equative copular sentences, this lends support to a predicational vs. equative ambiguity analysis of other languages (such as English) which only have one copula.

References

- Adger, David, and Gillian Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34:325-359.
- Barwise, Jon and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159-219.
- den Dikken, Marcel. 2006. *Relators and Linkers: the Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.
- Frank, Robert. 2002. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: MIT Press.
- Han, Chung-hye and Nancy Hedberg. 2008. Syntax and semantics of it-clefts: A Tree-Adjoining Grammar Analysis. *Journal of Semantics* 25:345-380.
- Heycock, Caroline and Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30:365-397.
- Higgins, F. Roger. 1973. The Pseudo-cleft Construction in English. PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA. Published 1979, Garland Press.
- Kuno, Susumo and Wongkhomthong, Preya. 1981. Characterizational and identificational sentences in Thai. *Studies in Language* 5:65-109.
- Mikkelsen, Line. 2005. *Copular Clauses: Specification, Predication and Equation*. Amsterdam/Philadelphia, PA: John Benjamins.
- Moro, Andrea. 1997. *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure*. Cambridge: Cambridge University Press.
- Partee, Barbara H. 1986. Ambiguous pseudoclefts with unambiguous *be*. *Proceedings of NELS 1986*. GLSA, University of Massachusetts, Amherst. 354-366.
- Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In J. Groenendijk and M. Stokhof (eds.), GRASS 8, *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Dordrecht: Foris.

Nancy Hedberg and David Potter

Williams, Edwin. 1983. Semantic versus syntactic categories. *Linguistics and Philosophy* 6:423-446.

Zaring, Laurie. 1996. Two "be" or not two "be": Identity, predication and the Welsh Copula. *Linguistics and Philosophy* 19:103-142.

Nancy Hedberg
Department of Linguistics
8888 University Drive.
Simon Fraser University
Burnaby, BC V5A 1S6
Canada

hedberg@sfu.ca

David Potter
Department of Linguistics
Northwestern University
2016 Sheridan Road
Evanston, Illinois 60208

dkp1@u.northwestern.edu