

Postnominal numerals are over-specified in referential communication: Evidence from Thai in comparison to Russian

© 2025

Natalia A. Zevakhina

HSE University, Moscow, Russia; Institute for Linguistic Studies, Russian Academy of Sciences,
St. Petersburg, Russia; natalia.zevakhina@gmail.com

Alina A. Shchipkova

HSE University, Moscow, Russia; schipkovalina@gmail.com

Alisa P. Chinkova

HSE University, Moscow, Russia; chinkova_a@mail.ru

Abstract: This paper presents experimental evidence for the over-specification (redundant use) of Thai postnominal numerals in referential communication. The first experiment reveals high rates of over-specification in oral responses that include Thai postnominal and Russian prenominal numerals tested in a contrastive multi-number visual context. The second experiment demonstrates lower but still relatively high rates of over-specification in written responses that include Thai postnominal numerals and color items tested in less contrastive two-number and bichrome visual contexts. Given this, the paper discusses the incrementality hypothesis and argues for the incremental production and processing of postnominal modifiers, which are explained by the visual salience of small cardinalities (2, 3, 4) and color and are regulated by syntactic linearization rules. The paper also provides additional evidence for the consistency which shapes the speaker's utterances as either over-specified or minimally specified throughout the whole communication.

Keywords: communication, experiment, numerals, redundancy, reference, Russian, Tai-Kadai, Thai

Acknowledgements: We express our sincere gratitude to Elena Pasalskaya for preparing the experimental materials, to Vera Ivanova for her invaluable help in glossing and translating the Thai data, and to Veronika Prigorkina and Vera Shumilina for their generous technical advice and support. We are grateful for the essential comments provided by two anonymous reviewers on the previous version of the manuscript. This study is an output of the research project No. 23-18-0069500695 supported by RSF, <https://rscf.ru/project/23-18-00695/>.

For citation: Zevakhina N. A., Schipkova A. A., Chinkova A. P. Postnominal numerals are over-specified in referential communication: Evidence from Thai in comparison to Russian. *Voprosy Jazykoznanija*, 2025, 2: 105–122.

DOI: 10.31857/0373-658X.2025.2.105-122

Избыточное употребление постномиальных числительных в референциальной коммуникации: данные тайского языка в сравнении с данными русского

Наталья Александровна Зевахина

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия;
Институт лингвистических исследований РАН, Санкт-Петербург, Россия; natalia.zevakhina@gmail.com

Алина Алексеевна Щипкова

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; schipkovalina@gmail.com

Алиса Павловна Чинкова

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; chinkova_a@mail.ru

Аннотация: В статье излагаются результаты эксперимента, посвященного избыточному употреблению тайских постноминальных числительных (т. е. числительных, которые располагаются после существительного) в референциальной коммуникации. В первом эксперименте выявлена высокая частотность избыточного употребления тайских постноминальных и русских преноминальных числительных в устных ответах испытуемых в контрастивном визуальном контексте с различным количеством тестируемых элементов. Во втором эксперименте обнаружена менее высокая частотность избыточного употребления тайских постноминальных числительных и цветообозначений в письменных ответах испытуемых в менее контрастивном контексте с двумя тестируемыми элементами двух цветов. В статье обсуждается гипотеза инкрементальности и выдвигается гипотеза пошагового порождения и обработки постноминальных модификаторов, которые объясняются визуальной выделенностью малых количеств (2, 3, 4) и цвета и регулируются правилами синтаксической линеаризации конкретного языка. В статье также приводится дополнительное свидетельство того, что говорящий на протяжении всей коммуникативной ситуации придерживается одной из двух стратегий: избыточного или минимального упоминания свойств предметов.

Ключевые слова: избыточность, коммуникация, референция, русский язык, тай-кадайские языки, тайский язык, числительные, эксперимент

Благодарности: Мы выражаем искреннюю благодарность Елене Пасальской за участие в подготовке экспериментального материала, Вере Ивановой за бесценную помощь в глоссировании и переводе тайских конструкций, Веронике Пригоркиной и Вере Шумиловой за грамотную подготовку документации. Исследование выполнено за счет гранта Российского научного фонда № 23-18-00695, <https://rscf.ru/project/23-18-00695/>.

Для цитирования: Zevakhina N. A., Schipkova A. A., Chinkova A. P. Postnominal numerals are over-specified in referential communication: Evidence from Thai in comparison to Russian. *Voprosy Jazykoznanija*, 2025, 2: 105–122.

DOI: 10.31857/0373-658X.2025.2.105-122

1. Introduction

The redundant specification of an object's properties in referential communication has extensively been investigated in experimental pragmatics. A typical research question is whether the speaker would label an object itself (e.g., “the shirt”) or an object and its property (e.g., “the green shirt”) in a situation when the speaker and listener are presented with several unique objects and the speaker directs the listener's visual search for a given pre-selected object. Naming an object is regarded as **minimal specification**, whereas additionally articulating its property (e.g., color) is viewed as **over-specification**.

Over-specification significantly varies depending on the properties under discussion and on the conditions under which it is tested. Although a considerable amount of experimental research on color, size, material, shape, pattern, location, and orientation has been conducted (cf. [Arts et al. 2011; Belke, Meyer 2002; Brown-Schmidt, Konopka 2011; Engelhardt et al. 2006; Gatt et al. 2013; Koolen et al. 2011; Long et al. 2021; Rubio-Fernandez 2016; 2019; 2021; Rubio-Fernandez et al. 2020; Tarenskeen et al. 2015; van Gompel et al. 2014], among others), color remains the most over-specified property. One of the reasons behind the high over-specification is its visual salience, which means that color is easily perceived. Another reason is that

color adjectives have absolute meanings, i.e., they are not context-dependent, unlike, e.g., size adjectives. However, color articulation is contingent upon the surrounding context. Recent studies (cf. [Belke, Meyer 2002; Koolen et al. 2013; Long et al. 2021; Rubio-Fernandez 2016; 2019; Rubio-Fernandez et al. 2020], among others) have demonstrated that, firstly, polychrome contexts boost color over-specification, whereas monochrome contexts decrease it. Secondly, the density of displayed items also plays a role: the denser the display, the higher the rates of color over-specification [Rubio-Fernandez 2019; 2021]. Thirdly, the semantic category a given object belongs to also influences the speaker's choice [Sedivy 2003; Westerbeek et al. 2015]. If a color is atypical or variable (rather than stereotypical) for a given semantic category, it is more likely to be articulated.

This empirical picture represents a major problem for Gricean pragmatics since it violates the second sub-maxim of the Gricean Quantity maxim: avoid redundancy. Indeed, why does the speaker excessively specify properties (e.g., color) of unique objects while referring to them in face-to-face communication? This question has been debated in the literature, see an overview in Section 1.1. In a nutshell, redundancy is a crucial component of referential communication and does not violate the Gricean principles of speaker-listener cooperation. On the contrary, redundancy makes communication optimal and efficient.

1.1. Incrementality, collaboration, efficiency

A detailed and influential approach to studying over-specification in referential communication is **Incremental Collaborative Efficiency (ICE)** model [Rubio-Fernandez 2021; Jara-Ettinger, Rubio-Fernandez 2022]; see also similar terms “incremental efficiency hypothesis” in [Rubio-Fernandez et al. 2020] and “efficient audience design” in [Long et al. 2021], among others. It consists of three components: incrementality, collaboration, and efficiency. Each of them should be distinguished from the other two.

Collaboration means that communication is a bipartite (speaker-centric and listener-centric) process. In a visual context identical for the interlocutors, color is a property mutually salient for both the speaker and listener. In other words, the speaker perceives some property as highly salient and believes that the listener perceives it similarly; therefore, the speaker articulates it. Likewise, the listener perceives the color as highly salient and believes that the speaker does so, too; therefore, the listener expects the speaker to produce it and easily recognizes it. This line of reasoning was inspired by the view of Clark and his colleagues, according to which referential communication is a collaborative process (cf. [Clark, Marshall 1981; Clark et al. 1983; Clark, Wilkes-Gibbs 1986; Clark, Schaefer 1989; Clark 1996]). It means that the speaker and listener accumulate common ground during conversation by making contributions regarded as a two-phase process: the speaker produces an utterance, and the listener accepts it (e.g., replies to it). Another idea that underlies the collaboration component of the ICE model is that communication is cooperative, as it is assumed in Gricean pragmatics [Grice 1975]. It means that interlocutors have joint goals to be achieved during their communication. According to Gricean pragmatics, they achieve their joint goals when the interlocutors' contributions are informative. However, Rubio-Fernandez and her colleagues have recently argued that it is not necessary. An utterance can be over-informative (i.e., it can convey more information than is needed for some current purposes) if it helps the interlocutors to achieve their joint goals. In other words, over-informative communication can be efficient.

Efficiency, which is the second component of the ICE model, means that over-specification can make referential communication easier than minimal specification if it facilitates the listener's search for an object in a visual context. This view is supported by Clark and Schaefer [1989] who argued that interlocutors aim to minimize their joint effort during their communication (see also [Rubio-Fernandez 2019]). This serves as an argument that although efficiency seems to contradict over-informativeness, it accords with the Gricean principle of cooperation.

The third component of the ICE model is **incrementality**. It means that messages are produced and processed unit by unit rather than globally (i.e., as whole utterances). This factor is determined by word order: the information conveyed by prenominal modifiers is articulated earlier by the speaker (and processed by the listener) than the information conveyed by postnominal modifiers. The idea has a direct psycholinguistic implication: it is assumed that prenominal adjectives are over-specified to a greater extent than postnominal adjectives. Rubio-Fernandez [2016; 2019], Rubio-Fernandez et al. [2020] convincingly demonstrated that English prenominal color adjectives are redundantly used significantly more often than Spanish postnominal color adjectives. According to Rubio-Fernandez et al. [2020], English listeners establish color contrast across categories (blue shapes > triangles), whereas Spanish listeners establish color contrast within a category (triangles > blue shapes). Wu and Gibson [2021] also observed the contrast between the two languages and provided further experimental evidence for incrementality in the use of numerals. The study focused only on small cardinalities 2, 3, and 4, but it did not provide any theory of why small cardinalities are over-specified. In their study, both English and Spanish prenominal numerals were used redundantly. This indicates that the contrast between the two languages lies in ordering rather than in native speakers' verbosity. A large-scale corpus analysis of 74 languages presented in [Kachakeche et al. 2021] provided additional evidence for incrementality. It showed that adjectival modification is more typical of languages with prenominal adjectives than of languages with postnominal adjectives. This is an interesting result, as cross-linguistically, postnominal adjectives exhibit a more frequent phenomenon than prenominal adjectives, according to the World Atlas of Language Structures [Dryer 2013b].

Another well-known approach to incrementality is **Continuous-Incremental Rational Speech Act** model, or CI-RSA model [Waldon, Degen 2021; Yu et al. 2023]. It is an extended model of the RSA model [Frank, Goodman 2012; Goodman, Frank 2016], which is a formalization of Gricean pragmatics [Grice 1975]. According to the CI-RSA model, utterances with prenominal adjectives (adjective + noun) are assigned higher probability than utterances with postnominal adjectives (noun + adjective). Language items produced earlier are predicted to have higher utility than language items produced later, which is conceptualized as a tradeoff between informativeness and cognitive cost.

Even though all the three components of successful referential communication have experimentally been proven to be key factors in over-specification, there seem to be some other factors that might interact. This idea is partially supported by some variation in prenominal and postnominal adjectives that is due to visual context manipulations, oral vs. written participant responses, the fluency or disfluency of responses, and some other factors (e.g., speakers' potential bilingualism). In the rest of this section, we focus on the first three factors and examine them one by one.

Firstly, Rubio-Fernandez [2016] pointed out that in polychrome visual contexts, the average rate of color over-specification is 95 % in English and only 59 % in Spanish. In monochrome visual contexts, the average rate of color over-specification is predictably lower: 37 % in English and 5 % in Spanish. Rubio-Fernandez et al. [2020] argued that the quantity of visual competitors also affects over-specification rates: a larger number of competitors boosts over-specification rates regardless of whether a language has prenominal or postnominal modification. This provides evidence for the influence of a visual context on the over-specification of properties.

Secondly, Wu and Gibson [2021] showed that the percentages of Spanish postnominal adjectives obtained in their paper were higher than the percentages received in [Rubio-Fernandez et al. 2020]: over 60 % vs. less than 10 % respectively. Crucially, their study was conducted in written form, which differs from the standard procedure of face-to-face communication experiments. Therefore, the task for the speakers was to identify a given picture and write down a description instead of orally interacting with the listener to help her identify a correct picture among others. This task implementation gave rise to an alternative production-based **incremental planning hypothesis** [Wu, Gibson 2021], see also [Pechmann 1989; Brown-Schmidt, Konopka 2008], which emphasizes a role of the utterance planning by the speaker in over-specified communication. It means that the speaker incrementally constructs her utterance regardless of whether she

addresses the listener directly. The speaker incrementally produces those features that she easily perceives in a visual context.

Thirdly, Brown-Schmidt and Konopka [2008] presented experimental evidence for the incremental planning of an utterance in the standard speaker-listener referential communication task, in which the interlocutors share an identical set of pictures. English and Spanish speakers produced adjectives at similar rates, resulting in fluent productions of postnominal size adjectives in Spanish (*la mariposa pequeña*, lit. ‘the butterfly small’) and disfluent productions of postnominal size adjectives in English (e.g., *the butterfly... uh small one*). This means that English speakers can use modifiers postnominally even though their utterances become disfluent.

We turn back to discussing incrementality, visual context manipulations, and collecting participant responses in Sections 1.3 and 2.

1.2. Consistency

Another parameter that has drawn attention in the literature on referential communication is consistency. As pointed out by Brennan and Clark [1996], interlocutors make a temporary agreement upon how they conceptualize a given object presented in a visual context by reusing lexical items. To illustrate, as the authors revealed, speakers might refer to a shoe as a *pennyloafer*, *casual shoe*, *brown shoe*, and *dress shoe*. In this regard, they adhere to some conceptualized and lexical strategy and expect their listeners to do the same. Pickering and Garrod [2004] argued for the interactive automatic alignment account that captures the fact that interlocutors align their representations at various linguistic levels in a dialogue. One of the consequences of this fact is lexical repetition in referential communication. Goodbeek and Krahmer [2012] established a connection between alignment and over-specification in referential communication. Interlocutors align their communicative behavior while interacting, i.e., they tend to include a redundant property in a referring expression if it was primed at earlier stages of the dialogue. Tarenskien et al. [2015] also observed that speakers are usually consistent with respect to how they refer to a given object throughout a whole dialogue. It means that they might use either a minimal specification of an object or its over-specification (e.g., redundantly specifying its properties). To illustrate, if the speaker describes a given object merely as a *hat*, she predictably does so during a whole experiment. Similarly, if the speaker describes a given object as a *red hat*, she is consistent in this respect. We call this the **consistency hypothesis**.

1.3. Small cardinalities, salience, and incrementality

In [Zevakhina et al. 2021] based on the data of Russian, we demonstrated that small cardinalities (2, 3, 4) represent yet another modifier type that is over-specified in referential communication due to their salience and absoluteness. The salience of small cardinalities was experimentally proven to be higher than the salience of color. Based on this finding, we formulate the **salience hypothesis** in the following way: the salience of small cardinalities is higher than the salience of color. The empirical fact that small cardinalities are over-specified is captured by the subitizing effect that is a human capacity to instantaneously grasp small cardinalities, that is, 2, 3, and 4 [Bourdon 1908; Kaufman et al. 1949]. Cardinalities more than 8 are large and involve effortful, error-prone, and slow computation, whereas cardinalities between 5 and 8 form an intermediate class. Their computation varies among people and is contingent upon the experimental setting (cf. [Jensen et al. 1950; Chi, Klahr 1975; Atkinson et al. 1976; Akin, Chase 1978; Oyama et al. 1981; Mandler, Shebo 1982]). According to Revkin et al. [2008], subitizing is a separate cognitive mechanism that is effortless, accurate, and rapid [Trick 1992].

It is well-acknowledged that numerals can have two meanings: ‘at-least’ meanings (‘at least n and possibly more’) and exact meanings (‘exactly n ’), cf. [Papafragou, Musolino 2003; Musolino 2004; Breheny 2008], among others. Zevakhina et al. [2021] argued that subitized small cardinalities are used in exact meanings and in this sense are absolute.

In [Zevakhina et al. 2021], we tested small cardinalities in a standard collaboration design, in which interlocutors share an identical visual context. The study revealed that they are mostly over-specified. In terms of the ICE model (see Section 1.1), the over-specification of small cardinalities demonstrates efficiency and collaboration. What has not been tested is the incremental nature of numerals denoting small cardinalities. Wu and Gibson [2021] argued that the over-specification of numerals that take prenominal position provides strong evidence for their incremental production. However, to the best of our knowledge, postnominal numerals have not been investigated. Relying upon the overview presented in Section 1.1, it seems natural to assume that postnominal numerals are over-specified to a lesser extent than prenominal ones. We refer to this as the **incrementality hypothesis**.

The rest of the paper is organized in the following way. Section 2 discusses the syntactic structures of the languages under consideration and enumerates the predictions of the present study verified in two experiments. Sections 3 and 4 report the methods and results of the first and second experiments, respectively. Section 5 concludes the paper.

2. Present study

The present study aims at investigating to what extent and how consistently small cardinalities are over-specified in Russian and Thai, which differ with respect to the prenominal vs. postnominal position of numerals. Since the language with postnominal numerals is also characterized with postnominal color items, the study additionally verifies the question whether postnominal positions of numerals and color items would be over-specified in a similar manner.

2.1. Word order in Thai and Russian

In what follows, we make an overview of the word order in Thai and Russian.

Thai numerals obligatorily occur after nouns (NP > NumP). Thai has classifiers that follow numerals. A linear syntactic structure is NP > NumP > ClassP, and a linear structure of syntactic heads is ‘noun + numeral + classifier’ (cf. [Noss 1964; Setthapun 1986; Smyth 2002; Iwasaki, Ingkaphirom 2009]). Thai nouns do not include plural markers. To illustrate, the phrase *three quadrangles* is translated as *siliem sam rup*, lit. ‘quadrangle three image’. The last word in this string *rup* ‘image’ is a classifier. The plural form without a numeral coincides with the singular form (e.g., *siliem* ‘quadrangle/quadrangles’). From a functional syntactic perspective, Thai has the following word orders: SVO, ‘noun + adjective’, and ‘noun + numeral + classifier’ (cf. [Warotamasikkhadit 1972], cited via [Dryer 2013a; 2013b; 2013c], respectively). From a generative grammar perspective, according to Singhapreecha [2001], who adopts the antisymmetry framework proposed by Kayne [1994], Thai numeral phrases are considered as syntactic complements of classifier phrases. Thai color items are also used in postnominal position. A linear structure is NP₁ > NP₂, where an NP₂ is presumably an adjunct of an NP₁ (e.g., *dao sideng*, lit. ‘star color red’). We thank the anonymous reviewer for pointing this out.

Russian has a free word order; however, the basic clause word order is SVO [Bivon 1971; Dryer 2013a]. The dominant word order in Russian numeral phrases is ‘numeral + noun’ [Bivon 1971; Dryer 2013c], as in many other Indo-European languages. According to Bailyn [2004], when a numeral takes the Nominative or inanimate Accusative case, it is in the Specifier position

of a quantifier phrase, and the Q head assigns the Genitive case to its complement, i.e., to a noun phrase (e.g., *tri kruga* ‘three.NOM circle.GEN.SG’). However, when a numeral takes any other morphological case assigned to it from a verb, it is in the Q head, and its complement, i.e., a noun phrase, is also assigned the same case by a verb (e.g., *dumat’o trëx krugax* ‘think about three.LOC circle.LOC.PL’). See also [Corbett 1993] for a theory-neutral analysis of Russian numerals. In referential communication, participants name objects and their cardinalities and therefore produce numerals in the Nominative case and nouns in the Genitive case.

Before moving on, let us make one remark. Could syntactic status and/or the (non-)obligatoriness of numerals and color items be a factor that might affect their over-specification? It is worth noting that, according to the cited analyses, Russian numerals represent either specifiers or heads of quantifier phrases, which select noun phrases as their complements [Bailyn 2004], whereas Thai numeral phrases are complements of classifiers [Singhapreecha 2001]. However, the “complex” ‘numeral + classifier’ in Thai is not obligatorily attached to noun phrases. To illustrate, the speaker might start her message with *siliem* ‘quadrangle/quadrangles’ and then either continue with *sam rup* ‘three image’ or not. In contrast, in Russian (and, to the best of our knowledge, in many other Indo-European languages), if the speaker produces a fluent utterance, she starts with a numeral and then proceeds with a noun phrase. For example, she might utter *tri* ‘three’, but then she needs to continue by saying the noun *kruga* ‘circle.GEN.SG’. The speaker might say it the other way round, starting with *krugi* ‘circle.NOM.PL’ and then disfluently adding *tri* ‘three’, but this is an unnatural way of delivering this message. Color items are more homogeneous with respect to (dis)fluency and (non-)obligatoriness across the languages considered here. In Russian, a color adjective usually takes a prenominal position but can be disfluently articulated in a postnominal position. In Thai, it is possible to say merely a noun (e.g., *dao* ‘star’) and then add a color item (*sideng* ‘red’), but it is not obligatory. In this respect, color constructions in Thai are similar to Spanish ones, in which the speaker starts with a noun and then can proceed with a color adjective, although it is not necessary. Therefore, Spanish color modifiers are adjuncts, and Thai ones can also be considered adjunct-like. Syntactic obligatoriness might facilitate over-specification, whereas syntactic non-obligatoriness might suppress it. Summing up, speakers can take into consideration the (non-)obligatoriness of numerals and color items while constructing their utterances. We refer to this as the **(non-)obligatoriness hypothesis**.

2.2. Predictions

The present study aims at verifying the following predictions.

The first group of predictions is: (i) small cardinalities are over-specified in referential communication, and (ii) their over-specification is determined by the position of numerals and their (non-)obligatoriness. Prediction (i) is motivated by the results obtained by Zevakhina et al. [2021], who argued that over-specification of small cardinalities is enabled due to their salience and absoluteness. Prediction (ii) is motivated, on the one hand, by the contrast between Spanish postnominal color adjectives vs. English prenominal color adjectives (cf. [Rubio-Fernandez 2016; 2019; Rubio-Fernandez et al. 2020]). Prediction (ii) accords with the incrementality hypothesis and can be reformulated as: the over-specification of Thai postnominal numerals is rarer than the over-specification of Russian prenominal numerals. On the other hand, prediction (ii) is motivated by a potential contrast between obligatory Russian vs. non-obligatory Thai numerals. Both predictions (i) and (ii) were verified in the first experiment (cf. Section 3).

Secondly, the study explores another prediction: (iii) consistency is observed in producing prenominal and postnominal numerals as well as in producing postnominal color items. It would accord with the findings received by Tarenskeen et al. [2015], who pointed out that the (over-)specification of attributes (such as color, pattern, and size) reveals consistency. If a speaker starts using some description (minimal specification or over-specification of attributes), she will stick

with it throughout the whole communication. Prediction (iii) was verified in both experiments (cf. Sections 3 and 4).

The third group of predictions is listed below. According to the incrementality hypothesis, (iv) both postnominal modifier types (color items and numerals) should be produced moderately. Prediction (v) says that Thai numerals are over-specified more often than color items. It is supported by the salience hypothesis: the over-specification of small cardinalities is more frequent than the over-specification of color due to the subitizing effect that gives rise to their greater degree of salience [Zevakhina et al. 2021]. Both predictions (iv) and (v) were verified in the second experiment (cf. Section 4).

Finally, even though the present study does not directly address the questions of how visual context manipulations and distinctions between oral vs. written responses affect over-specification rates, it is important to mention them here. The first experiment implemented a multi-number design (analogous to the polychrome visual contexts studied in the literature) to boost over-specification rates and make potential language differences clearer. The second experiment used a bi-chrome and two-number design that was implemented in the study by Zevakhina et al. [2021], see also [Tarenskeen et al. 2015]. The initial question was to see whether the latter design is also perceived as contrastive. However, there was an unforeseen obstacle. The first experiment of the present study collected oral responses, and the second experiment was also conceived to be conducted orally. For independent reasons, oral responses were impossible to collect, and therefore the second experiment collected written responses. Consequently, the two experiments differ in two factors and can only indirectly be compared to each other.

3. First experiment

3.1. Participants

In total, sixty-eight Thai and Russian native speakers voluntarily participated in the first experiment. All of them gave their consent to use their data anonymously for statistical purposes.

Thirty-three Thai native speakers participated in the Thai Condition. Their ages ranged from 21 to 34 years old, with the mean age of 26 years. Twenty-four were females, and nine were males.

Thirty-five Russian native speakers participated in the Russian Condition. Their ages varied from 20 to 40 years old, with the mean age of 23 years. Seventeen were females, and eighteen were males.

3.2. Methods

For the purposes of the experiment, we created 18 critical and 20 filler items which were slides with 2×2 cells. The cell to describe was highlighted.

Critical slides included uncolored geometric figures such as triangles, circles, squares, or diamonds. All figures within cells and across cells were of the same size. Each cell in a slide contained (1, 2, 3, or 4) identical uncolored geometric figures but the figures among all four cells in a slide were unique. Fig. 1a exemplifies a critical item.

We supposed that participants would use minimal specifications (figure) or over-specifications (figure + numeral + classifier) to describe highlighted objects. Importantly, the cardinality of 1 was never highlighted since it did not conform to the purposes of the experiment.

Filler items were 2×2 slides with human faces. Fig. 1b illustrates filler items used for both conditions.

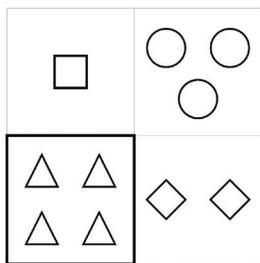


Fig. 1a. An example of a critical item used in the 1st experiment



Fig. 1b. An example of a filler item used in the 1st experiment

There were two practice items before the experiment.

Each participant saw only one experimental condition. The experiment was preceded by the instructions, according to which participants were presented with 40 slides with one highlighted cell to describe to their interlocutor, who had the same set of pictures, but the pictures in each slide were presented in a different order. The interlocutor did not know which of the pictures were highlighted. The task was to describe what was shown in the highlighted picture, so that the person would understand which picture it was. Due to independent circumstances, the first six participants in the Thai Condition were surveyed offline, whereas the others in the Thai Condition and all the participants in the Russian Condition were surveyed online. Both groups of participants saw an identical list of critical and filler items.¹

3.3. Results

In the Russian Condition, 630 responses were received. Among them, 2 responses (0.3%) were deleted since the figures were wrongly identified. In the rest of the 628 responses, both numeral and figure were over-specified in 574 responses (91%). This confirms prediction (i), that is, small cardinalities are over-specified in referential communication. In the Thai Condition, 594 responses were received. Among them, 10 responses (2%) were excluded because the figures were wrongly identified. In the rest of the 584 responses, both the numeral and the figure were over-specified in 561 responses (96%). This confirms prediction (i). In other words, both languages demonstrate the over-specification of small cardinalities.

The Wilcoxon rank sum test run for the results of the Russian and Thai Conditions demonstrated that Thai speakers over-specify small cardinalities more often than Russian speakers ($W = 174527$, $p < 0.001$), cf. Fig. 2 (p. 114). This does not support prediction (ii) and means that the over-specification of small cardinalities is not more typical of a language with prenominal numerals than of one with postnominal numerals.

Consistency was observed in both Russian and Thai responses. In the Russian Condition, three participants (8.6%) provided minimal descriptions in most (approximately 83%) or all of their responses. Among the others, twenty-nine participants (approximately 83%) provided over-specified descriptions only. This confirms prediction (iii). In the Thai Condition, one participant (3%) minimally described all the pictures throughout the experiment. Twenty-nine participants (approximately 88%) over-informatively described most of the pictures (at least 94%), whereas three participants (9%) minimally specified only the first critical item and over-specified

¹ The materials for both experiments can be accessed at <https://osf.io/dhztq/>. All the responses for Thai and Russian Conditions were glossed and categorized either as minimal specifications or over-specifications.

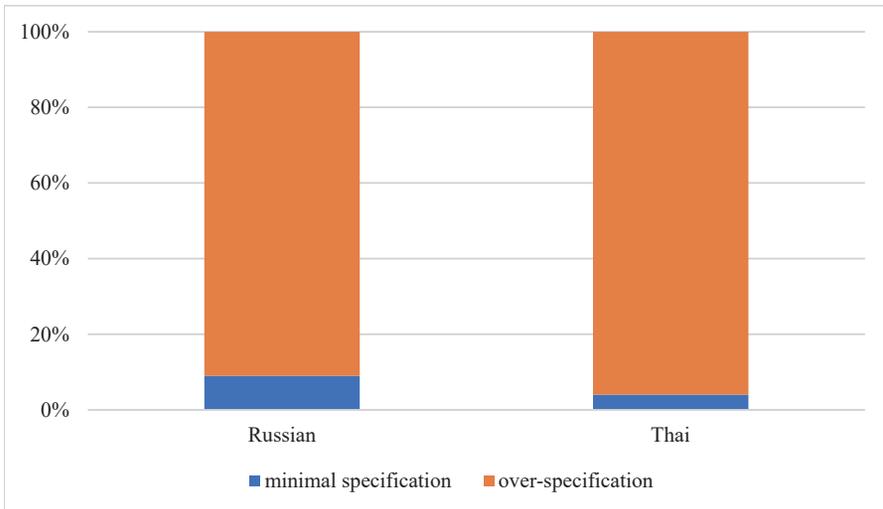


Fig. 2. Distribution of over-specification vs. minimal specification in the Thai and Russian Conditions

the rest. One participant (3 %) minimally specified only the first two critical items and over-specified the rest. These results support prediction (iii). The results of both Russian and Thai Conditions suggest that the participants chose one of the two possible strategies: either to over-specify cardinality throughout all their responses or not.

Finally, we considered the distribution of Thai classifiers among the responses. The classifier *an* ‘thing’ was used in 201 responses (34 %), the classifier *rup* ‘image’ can be found in 262 responses (45 %), and the classifier *vong* ‘round object’ is attested in 69 responses (12 %), the rest were the classifiers *tua* ‘figure’, *chin* ‘part’ and some of the combinations, e.g., *rup vongklom song an* ‘image circle two thing’, *rup samliem si an* ‘image triangle four thing’, literal translations.

3.4. Discussion

The results of the first experiment showed that small cardinalities are over-specified in a contrastive, multi-number design in both languages, even though they are genealogically unrelated and structurally different. They sustained the findings by Zevakhina et al. [2021]. The tendency to over-specify was consistent, and this is in line with the data presented in [Tarenskeen et al. 2015]. Therefore, the consistency hypothesis was corroborated. The experiment revealed a remarkable finding: the postnominal and prenominal positions of numerals provided over-specified descriptions more than 90 % of the time, and postnominal descriptions in Thai were more frequent than prenominal ones in Russian. These results do not support the incrementality hypothesis as formulated in Section 1.3, and therefore contrast with the results reported for the prenominal vs. postnominal position of color adjectives in the papers by Rubio-Fernandez and her colleagues.

One plausible reason for this is that small cardinalities possess a high degree of salience, necessitating speakers to report them and to produce numerals even in postnominal position. This sort of salience is explained by the subitizing effect observed in small cardinalities. It is important to stress here that speakers consider that their interlocutors also perceive the visual salience of small cardinalities, and therefore, when they produce numerals in postnominal position, they facilitate listeners’ visual search for a given object. In this sense, referential communication is

collaborative and efficient and accords with the ICE model. An implication for the distinction between prenominal vs. postnominal modifiers is that the articulation of modifiers is regulated by the syntactic linearization of a given language. In both languages considered in the present study, a high degree of visual perceptibility of small cardinalities is reflected in the choice of referential expressions (i.e., numerals). However, the visual salience of properties and their lexical choice are restricted by syntactic structure, resulting in postposition or preposition.

Syntactic (non-)obligatoriness as another plausible syntactic reason for the observed distinction between the two types of numerals does not hold. Russian numerals are typically located in the prenominal position, determine the noun's morphological form and are obligatory, whereas Thai numeral phrases take a postnominal position and are non-obligatory. However, as the data showed, obligatoriness does not increase over-specification rates. Therefore, the (non-)obligatoriness hypothesis is not confirmed.

To test the idea that small cardinalities are over-specified to a greater extent than color in postnominal position, we conducted a second experiment. We made the visual displays less contrastive (two-number for small cardinalities and bichrome for color) to see whether over-specification rates would remain high.

4. Second experiment

4.1. Participants

In total, seventy-six Thai native speakers voluntarily participated in the second experiment. All of them gave their consent to use their data anonymously for statistical purposes.

Forty-four of them replied to the Color Condition. Two of them were underage (14 and 16 years old); therefore, we did not consider their data. This resulted in forty-two participants for the Color Condition aged between 18 and 55, with a mean age of 32 years, thirty-two females and ten males.

Thirty-two of others answered the Small Cardinality Condition. Their ages varied from 19 to 40 years, and the mean age was 22 years, twenty females and twelve males.

4.2. Methods

Although the design of the second experiment was in general analogous to that of the first experiment, some modifications were made.

For independent reasons, it was impossible to interview the participants *tête-à-tête*, even online via Skype or Zoom. As a result, the experiment was transformed into a Google Forms survey to collect participants' written responses (see [Wu, Gibson 2021] and Section 1.1 for the method of data collection in written form).

For experimental purposes, two Google Forms were created for the Cardinality and Color Conditions. The materials were similar to the ones used in the first experiment in [Zevakhina et al. 2021]. Each condition was supplied with 60 (30 critical and 30 filler) slides that included 2 x 2 cells. The cell to describe was highlighted. Critical slides included geometric figures such as circles, ovals, squares, rectangles, triangles, diamonds, stars, and pluses. Thus, the second experiment had more slides and more geometric figures than the first one. All figures within and across the cells were of the same size.

In the Small Cardinality Condition, each cell in a slide contained (2, 3, or 4) identical uncolored geometric figures, but the figures across all four cells in a slide were unique. We removed

cardinal 1 since it was never highlighted in the first experiment, and it seemed to be redundant. The two cells placed diagonally included an identical number of figures and differed from two other same-number cells which were also placed diagonally. This two-number design differed from the multi-number design used in the first experiment. The reason behind this was that the multi-number design might boost over-specification, and therefore we intended to investigate whether the tendency toward the over-specification of small cardinalities would be preserved in a two-number design. In total, there were 30 critical slides in this condition. Fig. 3a exemplifies a critical item for the Small Cardinality Condition.

In the Color Condition, one cell included only one (red, green, or yellow) colored geometric figure. In each slide, two cells placed diagonally were identically colored and differed from two other same-color cells placed diagonally. In total, there were 30 critical slides in this condition. Fig. 3b illustrates a critical item for the Color Condition.

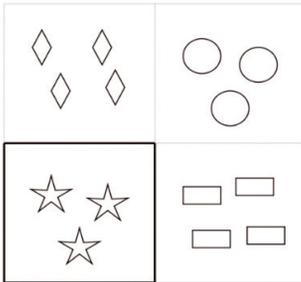


Fig. 3a. Example of a critical item for the Small Cardinality Condition used in the 2nd experiment

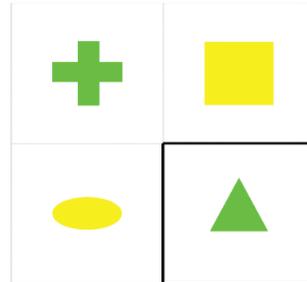


Fig. 3b. Example of a critical item for the Color Condition used in the 2nd experiment

We supposed that participants would use minimal specifications (e.g., ‘star’) or over-specifications (e.g., ‘star three’ + classifier) to describe highlighted objects in the Small Cardinality Condition. We also expected participants to use minimal specifications (e.g., ‘triangle’) or over-specifications (e.g., ‘triangle green’) to describe the highlighted objects in the Color Condition.

Filler items were 2 x 2 slides either with human faces, tangrams, or various objects (clothes, transport, furniture). In total, there were 30 filler items, which were identical for both conditions. Fillers were more diverse in the second experiment than in the first one. We supposed that descriptions of objects would be relatively easy, descriptions of tangrams would be difficult, and descriptions of human faces would be in-between. Fig. 4a–4b illustrate filler items used for both conditions.

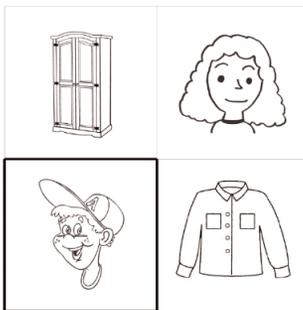


Fig. 4a. Example of a filler item used in the 2nd experiment (human faces and various objects)

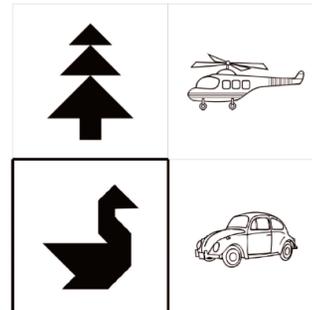


Fig. 4b. Example of a filler item used in the 2nd experiment (tangrams and transport)

Each participant received one out of two links to the Google Forms and saw only one experimental condition. The experiment was preceded with the instructions in Thai, according to which participants would be presented with 60 slides with one highlighted cell to describe by writing a description under the slide. The slides had to be described in Thai.

4.3. Results

In total, 1260 responses (42 participants x 30 critical items) were received for the Color Condition, and 960 responses (32 participants x 30 critical items) were obtained for the Small Cardinality Condition. Only Thai responses were considered (English responses were excluded). As a result, 30 responses by one participant in the Small Cardinality Condition, 30 responses by one participant in the Color Condition and one response by another participant were excluded.

Several other cases were ruled out. Eight under-informative responses that occurred in the Color Condition were also excluded (e.g., several answers included a color item without a figure name). Secondly, ten wrong responses that occurred in the Cardinality and Color Conditions were excluded. They either contained a wrong figure identification, a wrong quantity identification or lacked a classifier. For example, in a situation where four rectangles had to be described, a response ‘triangle’ was given; in a situation where two stars had to be described, a response ‘star four things’ was provided; in a situation where three rectangles had to be described, a response ‘rectangle four’ was given. Thirdly, metaphorical figure names were not considered. To illustrate, a diamond might be referred to as ‘a sweet’ or an oval could be referred to as ‘an egg’. Five metaphorical names occurred in the Small Cardinality Condition, and 18 metaphorical names occurred in the Color Condition. Fourthly, four responses from the first four participants were not collected due to a technical error in the Google Form: no text field appeared. Later, this error was corrected, and other participants completed the whole task. Four responses were not intentionally given by one participant in the Color Condition: a dash was inserted instead, although other slides with the same figure were correctly described.

Consequently, 1198 responses were given for the Color Condition, and 915 responses were provided for the Small Cardinality Condition. All the responses were glossed and categorized either as minimal specifications or as over-specifications. Recall that an expression is over-specified if it includes a color item or a numeral with a classifier.

The answers in the Color Condition divide into two categories: (i) minimally informative answers that included only a figure name (e.g., *samliem* ‘triangle’, *vongklom* ‘circle’) and (ii) over-informative answers that comprised a figure name and a color item (e.g., *samliem sikhiao* ‘triangle color green’, *buak siluang* ‘plus color yellow’, *dao sideng* ‘star red’, literal translations).

The answers in the Small Cardinality Condition were more diverse: (i) minimally informative answers that included only figure names (examples are identical to the ones provided for (i) in the Color Condition); (ii) over-informative answers that included a figure name, a numeral, and a classifier (e.g., *samliem si an* ‘triangle four things’, *siliem sam rup* ‘quadrangle three images’, *vongklom song vong* ‘circle two round things’, *vong ri song vong* ‘oval two round things’, *buak song chin* ‘plus two parts’, *dao si duang* ‘star four celestial body’, *khaolam tad si tua* ‘diamond four bodies’, literal translations); (iii) over-informative answers that comprised a word for an image, a figure name, a numeral, and a classifier (e.g., *rup vongklom song vong* ‘image circle two round things’, *rup samliem sam rup* ‘image triangle three images’, *rup buak si an* ‘image plus four things’, literal translations).

The Color Condition comprised 841 (out of 1198) over-specified responses (approximately 70%), whereas the Small Cardinality Condition included 621 (out of 912) over-specified responses (approximately 68%), see Fig. 5. Among the over-specified responses in the Small Cardinality condition, 585 included a figure name, a numeral, and a classifier (cf. group (ii) in the

previous paragraph), whereas 36 comprised a word for an image, a figure name, a numeral, and a classifier (cf. group (iii) in the previous paragraph), approximately 64 % vs. 4 % respectively.

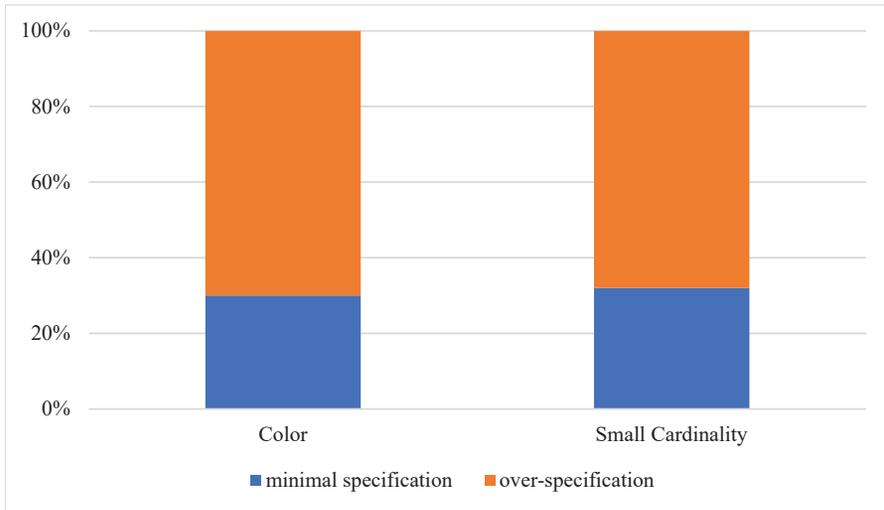


Fig. 5. Distribution of minimally specified and over-specified responses in the Color and Small Cardinality Conditions

The data confirm prediction (iv), according to which postnominal color and cardinality modifiers were expected to be over-specified moderately. The Wilcoxon rank sum test demonstrated that the distribution of minimally specified and over-specified responses for color and cardinality items was similar ($W = 534771$, $p = 0.2985$). This does not support prediction (v) that Thai numerals are over-specified more often than color items.

In the Small Cardinality Condition, 18 participants (approximately 56 %) provided over-specified responses to at least 24 critical items (80 %). Five participants (15.6 %) provided minimally specified responses to at most one critical item (approximately 3 %). In the Color Condition, 26 participants (59 %) submitted over-specified responses to at least 24 critical items (80 %). Nine participants (20.45 %) submitted minimally specified responses to no critical item. This suggests that prediction (iii) was confirmed: consistency is observed in producing postnominal numerals and postnominal color items.

In all the responses of the Small Cardinality Condition, the numerals *song*, *si*, and *sam* 2, 3, and 4 were used with the following classifiers. The classifier *an* ‘thing’ (49.68 %) was the most frequent throughout the experiment and could be used with almost any noun. The classifier *rup* ‘image’ (15.22 %) is relevant since it describes pictures or photos. The classifier *vong* ‘round object’ (14.9 %) is used to describe round objects and was used to describe circles and ovals in the present experiment. The classifier *chin* ‘part’ (8.81 %) was rarer. The classifier *duang* ‘celestial body’ (9.62 %) was used only in the descriptions of stars. The rest were classifiers *loom* ‘hole’, *kruangmai* ‘sign, symbol’, and *tua* ‘body’.

4.4. Discussion

The results of the second experiment corroborated the idea that postnominal modifiers (color items and numerals) are produced moderately. The findings demonstrated that the tendency to over-specify properties is consistent, and this is in line with the data of the first experiment

and with the data reported in [Tarenskeen et al. 2015]. However, the results did not confirm the expectation that Thai numerals are over-specified more often than color items. In what follows, we consider these results and their implications.

Firstly, Thai revealed moderate proportions of the specification of both color and small cardinalities. As far as we can see, there are two reasons for this.

One reason is that participants' responses were collected in written form. Comparing the data of the first experiment with the data of the second one, we see that the over-specification proportions received in oral responses were higher than the ones received in written responses. This fact contradicts the empirical picture presented in the papers by Rubio-Fernandez et al. [2020] and Wu and Gibson [2021]. In their studies, the over-specification rates obtained in oral responses were lower than the ones obtained in written responses. However, the rates for written responses in our second experiment were still relatively high (more than 65 %), presumably because color and small cardinalities are visually salient and therefore are straightforwardly recognizable. In this respect, they support the incremental planning hypothesis. The rates for oral responses in our first experiment were even higher (more than 90 %), providing evidence for the fact that oral face-to-face communication increases over-specification rates in contrast to written production. This confirms the collaboration and efficiency components of the ICE model. To make communication collaborative and efficient, the speaker redundantly specifies properties of a given object to help the listener identify it in a visual context.

Another reason for why the data of the second experiment were moderate in comparison to the data of the first experiment is that the design of the former was less contrastive (bichrome and two-number) than that of the latter (multi-number). This is an expected finding since it accords with the results in [Tarenskeen et al. 2015]: in contrastive two-color, pattern, and size displays, the over-specification rates were not higher than 80 %. Furthermore, it partially accords with the fact that visual contrast is an important factor in over-specification rates: increasing rates for polychrome displays, which are highly contrastive, vs. decreasing rates for monochrome displays, which are less contrastive, see [Belke, Meyer 2002; Koolen et al. 2013; Long et al. 2021; Rubio-Fernandez 2016; 2019; Rubio-Fernandez et al. 2020], among others.

Be that as it may, since the responses were collected in written form, we cannot say for sure which factor (the way of collecting responses or the experimental design) played a more important role in the moderate percentages.

Secondly, the results demonstrated that the over-specification rates for color and small cardinalities were similar in a language with postnominal modifiers. The results obtained by Zevakhina et al. [2021] revealed that small cardinalities were over-specified more often than color in Russian, which is a language with prenominal modifiers. Does the reason lie in the modifier position? It does not look plausible since several papers [Rubio-Fernandez 2016; 2019; Rubio-Fernandez et al. 2020; Wu, Gibson 2021] pointed out that prenominal modifiers are more likely to be over-specified than postnominal modifiers. Therefore, for the time being, it is hard to conclude which parameter, color or small cardinality, is more salient, since presumably other factors also play a role. A tentative conclusion would be that both parameters are visually salient and therefore are to be communicated. Consequently, the salience hypothesis formulated in Section 1.3 should be modified. A new version of the salience hypothesis is that both color and small cardinalities are salient and yield over-specified descriptions in referential communication.

Thirdly, in both our experiments postnominal modifiers were over-specified in more than 60 % of the utterances. Even though it is not compatible with the previous experimental data for English and Spanish [Rubio-Fernandez 2016; 2019; Rubio-Fernandez et al. 2020; Wu, Gibson 2021] and with the version of the incrementality hypothesis formulated in Section 1.1, we believe that our findings do not contradict the incrementality mechanism observed in natural languages; rather, the hypothesis should be amended. Given the data, we argue that the speaker incrementally constructs, and the hearer incrementally comprehends utterances in referential communication, on the one hand, in accordance with the visual salience of objects and their properties and,

on the other hand, in accordance with syntactic linearization rules which impose word order restrictions in a given language. Therefore, syntactic linearization represents an important factor affecting whether and how Thai speakers include numerals and color items in their descriptions.

5. Conclusion

The two experiments reported in this paper demonstrated that postnominal numerals and postnominal color items in Thai are typically over-specified in referential communication. This is due to the visual salience of small cardinalities and color that is easily perceived by both interlocutors, and each of them is aware that the other interlocutor also perceives these properties. The speaker produces a property and believes that this production makes the listener's search for an object more efficient and easier. Speakers describe objects in a consistent manner either by over-specifying or minimally specifying their properties. The report of a salient property is regulated by the syntactic linearization of a given language (whether modifiers take prenominal or postnominal positions). This syntactic mechanism forces speakers to articulate an utterance according to the word order of a given language. This means that an utterance is produced incrementally regarding visual perceptibility and syntactic rules.

ABBREVIATIONS

GEN — genitive case
LOC — locative case

NOM — nominative case
SG — singular

PL — plural

REFERENCES

- Akin, Chase 1978 — Akin O., Chase W. Classification of three-dimensional structures. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4(3): 397–410.
- Arts et al. 2011 — Arts A., Maes A., Noordman L., Jansen C. Overspecification in written instruction. *Linguistics*, 2011, 49(3): 555–574.
- Atkinson et al 1976 — Atkinson J., Campbell F. W., Francis M. R. The magic number 4 ± 0 : A new look at visual numerosity judgements. *Perception*, 1976, 5(3): 327–334.
- Bailyn 2004 — Bailyn J. The case of Q. *Annual Workshop on Formal Approaches to Slavic Linguistics. The Ottawa Meeting 2003*. Arnaudova O., Browne W., Rivero M. L., Stojanović D. (eds.). Ann Arbor: Michigan Slavic Publ., 2004, 1–35.
- Belke, Meyer 2002 — Belke E., Meyer A. Tracking the time course of multidimensional stimulus discrimination: Analyses of viewing patterns and processing times during “same”–“different” decisions. *European Journal of Cognitive Psychology*, 2002, 14(2): 237–266.
- Bivon 1971 — Bivon R. *Element order*. Cambridge: Cambridge Univ. Press, 1971.
- Bourdon 1908 — Bourdon B. Sur le temps nécessaire pour nommer les nombres. *Revue Philosophique de la France et de l'Étranger*, 1908, 65: 426–431.
- Breheny 2008 — Breheny R. A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics*, 2008, 25: 93–139.
- Brennan, Clark 1996 — Brennan S., Clark H. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1996, 22(6): 1482–1493.
- Brown-Schmidt, Konopka 2008 — Brown-Schmidt S., Konopka A. E. Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, 2008, 109(2): 274–280.
- Brown-Schmidt, Konopka 2011 — Brown-Schmidt S., Konopka A. E. Experimental approaches to referential domains and the on-line processing of referring expressions in unscripted conversations. *Information*, 2011, 2(2): 302–326.

- Chi, Klahr 1975 — Chi M. T. H., Klahr D. Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, 1975, 19(3): 434–439.
- Clark 1996 — Clark H. H. *Using language*. Cambridge: Cambridge Univ. Press, 1996.
- Clark, Marshall 1981 — Clark H. H., Marshall C. R. Definite reference and mutual knowledge. *Elements of discourse understanding*. Joshi A. K., Webber B., Sag I. A. (eds.). Cambridge: Cambridge Univ. Press, 1981, 10–63.
- Clark, Schaefer 1989 — Clark H. H., Schaefer E. F. Contributing to discourse. *Cognitive Science*, 1989, 13(2): 259–294.
- Clark et al. 1983 — Clark H. H., Schreuder R., Buttrick S. Common ground at the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 1983, 22(2): 245–258.
- Clark, Wilkes-Gibbs 1986 — Clark H. H., Wilkes-Gibbs D. Referring as a collaborative process. *Cognition*, 1986, 22(1): 1–39.
- Corbett 1993 — Corbett G. G. The head of Russian numeral expressions. *Heads in grammatical theory*. Corbett G. G., Fraser N. M., McGlashan S. (eds.). Cambridge: Cambridge Univ. Press, 1993, 11–35.
- Dryer 2013a — Dryer M. Order of subject, object and verb. *WALS Online*. Dryer M., Haspelmath M. (eds.). <http://wals.info/chapter/81>.
- Dryer 2013b — Dryer M. Order of adjective and noun. *WALS Online*. Dryer M., Haspelmath M. (eds.). <http://wals.info/chapter/87>.
- Dryer 2013c — Dryer M. Order of numeral and noun. *WALS Online*. Dryer M., Haspelmath M. (eds.). <http://wals.info/chapter/89>.
- Engelhardt et al. 2006 — Engelhardt P. E., Bailey K. G. D., Ferreira F. Do speakers and listeners observe the Gricean maxim of quantity? *Journal of Memory and Language*, 2006, 54(4): 554–573.
- Frank, Goodman 2012 — Frank M. C., Goodman N. D. Predicting pragmatic reasoning in language games. *Science*, 2012, 336(6084), 998.
- Gatt et al. 2013 — Gatt A., Krahmer E., van Gompel R. P. G., van Deemter K. Production of referring expressions: Preference trumps discrimination. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 2013: 483–488.
- Goodman, Frank 2016 — Goodman N. D., Frank M. C. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 2016, 20(11): 818–829.
- Goudbeek, Krahmer 2012 — Goudbeek M., Krahmer E. Alignment in interactive reference production: Content planning, modifier ordering, and referential overspecification. *Topics in Cognitive Science*, 2012, 4(2): 269–289.
- Grice 1975 — Grice H. P. Logic and conversation. *Syntax and semantics*. Vol. 3: *Speech acts*. Cole P., Morgan J. L. (eds.). New York: Academic Press, 1975, 41–58.
- Iwasaki, Ingkaphirom 2009 — Iwasaki S., Ingkaphirom P. *A reference grammar of Thai*. Cambridge: Cambridge Univ. Press, 2009.
- Jara-Ettinger, Rubio-Fernandez 2022 — Jara-Ettinger J., Rubio-Fernandez P. The social basis of referential communication: Speakers construct physical reference based on listeners' expected visual search. *Psychological Review*, 2022, 129(6): 1394–1413.
- Jensen et al. 1950 — Jensen E. M., Reese E. P., Reese T. W. The subitizing and counting of visually presented fields of dots. *Journal of Psychology*, 1950, 30: 363–392.
- Kachakeche et al. 2021 — Kachakeche Z., Futrell R., Scontras G. Word order affects the frequency of adjective use across languages. *Proceedings of the 43th Annual Meeting of the Cognitive Science Society*, 2021: 3006–3012.
- Kaufman et al. 1949 — Kaufman E. L., Lord M. W., Reese T. W., Volkman J. The discrimination of visual number. *The American Journal of Psychology*, 1949, 62(4): 498–525.
- Kayne 1994 — Kayne R. S. *The antisymmetry of syntax*. Cambridge (MA): MIT Press, 1994.
- Koolen et al. 2011 — Koolen R., Gatt A., Goudbeek M., Krahmer E. Factors causing overspecification in definite descriptions. *Journal of Pragmatics*, 2011, 43(13): 3231–3250.
- Koolen et al. 2013 — Koolen R., Goudbeek M., Krahmer E. The effect of scene variation on the redundant use of color in definite reference. *Cognitive Science*, 2013, 37(2): 395–411.
- Long et al. 2021 — Long M., Moore I., Mollica F., Rubio-Fernandez P. Contrast perception as a visual heuristic in the formulation of referential expressions. *Cognition*, 2021, 217, 104879.
- Mandler, Shebo 1982 — Mandler G., Shebo B. Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, 1982, 111(1): 1–22.
- Musolino 2004 — Musolino J. The semantics and acquisition of number words: Integrating linguistic and developmental perspectives. *Cognition*, 2004, 93(1): 1–41.

- Noss 1964 — Noss R. B. *Thai reference grammar*. Washington, DC: Foreign Service Institute, 1964.
- Oyama et al. 1981 — Oyama T., Kikuchi T., Ichihara Sh. Span of attention, backward masking and reaction time. *Perception and Psychophysics*, 1981, 29(2): 106–112.
- Papafragou, Musolino 2003 — Papafragou A., Musolino J. Scalar implicatures: Experiments at the syntax-semantics interface. *Cognition*, 2003, 86(3): 253–282.
- Pechmann 1989 — Pechmann T. Incremental speech production and referential overspecification. *Linguistics*, 1989, 27(1): 89–110.
- Pickering, Garrod 2004 — Pickering M. J., Garrod S. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Science*, 2004, 27: 169–226.
- Revkin et al. 2008 — Revkin S. K., Piazza M., Izard V., Cohen L., Dehaene S. Does subitizing reflect numerical estimation? *Psychological Science*, 2008, 19(6): 607–614.
- Rubio-Fernandez 2016 — Rubio-Fernandez P. How redundant are redundant color adjectives? An efficiency-based analysis of color overspecification. *Frontiers in Psychology*, 2016, 7, 153.
- Rubio-Fernandez 2019 — Rubio-Fernandez P. Over-informative speakers are cooperative: Revisiting the Gricean maxim of quantity. *Cognitive Science*, 2019, 43(11), e12797.
- Rubio-Fernandez 2021 — Rubio-Fernandez P. Color discriminability makes over-specification efficient: Theoretical analysis and empirical evidence. *Humanities and Social Sciences Communications*, 2021, 8, 147.
- Rubio-Fernandez et al. 2020 — Rubio-Fernandez P., Mollica F., Jara-Ettinger J. Speakers and listeners exploit word order for communicative efficiency: A cross-linguistic investigation. *Journal of Experimental Psychology: General*, 2020, 150(3): 583–594.
- Sedivy 2003 — Sedivy, J. C. Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, 2003, 32: 3–23.
- Setthapun 1986 — Setthapun W. *Simple Thai grammar. Characteristics of the Thai language, pronunciation guide and the essential aspects of grammar*. London: Europhone Language Centre, 1986.
- Singhapreecha 2001 — Singhapreecha P. Thai classifiers and the structure of complex Thai nominals. *Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation*, 2001: 259–270.
- Smyth 2002 — Smyth D. *Thai: An essential grammar*. London: Routledge, 2002.
- Tarenskeen et al. 2015 — Tarenskeen S., Broersma M., Geurts B. Overspecification of color, pattern, and size: saliency, absoluteness, and consistency. *Frontiers in Psychology*, 2015, 6, 1703.
- Trick 1992 — Trick L. A theory of enumeration that grows out of a general theory of vision: Subitizing, counting, and FINSTs. *The nature and origins of mathematical skills*. Campbell J. (ed.). Amsterdam: Elsevier, 1992, 257–299.
- van Gompel et al. 2014 — van Gompel R. P., Gatt A., Krahmer E., van Deemter K. Overspecification in reference: Modelling size contrast effects. Poster presented at *Architectures and Mechanisms for Language Processing, Edinburgh, 2014*.
- Waldon, Degen 2021 — Waldon B., Degen J. Modeling cross-linguistic production of referring expressions. *Proceedings of the Society for Computation in Linguistics*, 2021, 4: 206–215.
- Warotamasikkhadit 1972 — Warotamasikkhadit U. *Thai syntax*. The Hague: Mouton, 1972.
- Wu, Gibson 2021 — Wu S., Gibson E. Word order predicts cross-linguistic differences in the production of redundant color and number modifiers. *Cognitive Science*, 2021, 45(1), e12934.
- Westerbeek et al. 2015 — Westerbeek H., Koolen R., Maes A. Stored object knowledge and the production of referring expressions: The case of color typicality. *Frontiers in Psychology*, 2015, 6, 935.
- Yu et al. 2023 — Yu D., Waldon B., Degen J. The cross-linguistic order of adjectives and nouns may be the result of iterated pragmatic pressures on referential communication. *Proceedings of the 45th Annual Meeting of the Cognitive Science Society*, 2023: 1003–1010.
- Zevakhina et al. 2021 — Zevakhina N., Pasalskaya E., Chinkova A. Over-specification of small cardinalities in referential communication. *Frontiers in Psychology*, 2021, 12, 745230.