CONVENTIONAL AND UNCONVENTIONAL METONYMY: CAN YOU CURL UP WITH A GOOD AGATHA CHRISTIE IN YOUR SECOND LANGUAGE?

Roumyana Slabakova[1,2]

Jennifer Cabrelli Amaro[3]

Sang Kyun Kang[2]

[1]University of Southampton, [2]University of Iowa, [3]U. S. Air Force Academy

Abstract

This article presents results of two off-line comprehension tasks investigating the acceptability of unconventional and conventional metonymy by native speakers of Korean and Spanish who speak English as a second language. We are interested in discovering whether learners differentiate between conventional and unconventional metonymy, and whether the acceptability of metonymic expressions in the native language has an effect on learners' judgments in the second language. The findings of this study constitute further experimental support for the psychological reality of the distinction between conventional and unconventional metonymy, but only in English. Learners of English at intermediate levels of proficiency exhibit transfer from the native language in comprehending metonymic shifts of meanings. Restructuring of the grammar is evident in later stages of development. Finally, complete success in acquiring L2 metonymic patterns is attested in our experimental study. Implications for L2A theories and teaching practices are discussed.

Keywords: metonymy comprehension, conventional (regular) metonymy, unconventional metonymy, Korean, Spanish, English, Bottleneck Hypothesis

INTRODUCTION

Human language offers numerous examples of non-compositional utterances, where the meaning is more than the sum of its parts. Take for example the utterances below from recent and classic sources. Example (1a) is taken from the recent movie *Silver Linings Playbook*. It is produced in the following context: Pat and Tiffany go on a date in a diner. Pat orders Raisin Bran because he doesn't want the occasion to seem like a date, and she orders tea. The waitress brings the Raisin Bran (with milk) and the tea over to their table. The characters quarrel about something and Tiffany storms out of the diner. Pat tries to follow her out when the waitress stops him and says:

(1)   a. Slow down, <u>Raisin Bran</u>!

         [*meaning*: the person associated with what he ordered]

    b. <u>Ringo</u> was hit in the fender by a truck when he was momentarily distracted by a

         motorcycle.   (Nunberg 1995, modified from Jackendoff 1992)

         [*meaning*: the car Ringo was driving]

    c. [One flight attendant to another]:

         Ask <u>seat 19</u> whether he wants to swap. (Markert and Nissim 2006: 159, British

         National Corpus)

         [*meaning*: the person sitting in seat 19]

All the utterances exhibit *additional meaning* beyond the meanings of its words. Thus this is a special case of polysemy (the coexistence of two or more meanings attributed to a word or phrase) (Kövecses & Radden 1998). The *literal meaning* of the underlined phrase, determined by the meanings of the constituent words, is enriched to a larger, pragmatically-determined

meaning given in square brackets under the examples. Such examples are well known and discussed in the literature as *transfer of reference* (Nunberg 1979, 1995), *enriched composition* (Jackendoff 1997), or *metonymy* (e.g. Lakoff 1987).

Metonymy is a well-established mental process, whereby the mention of some entity (activity, person, thing, time period, etc.) is interpreted *to stand for* a related entity. Typical substitutions of this sort include the capital for the country's government as in (2a), a place name for an event as in (2b), and a place name for a typical product (2c).

(2)   a.   But it was unclear whether <u>Beijing</u> would meet past UN demands for unrestricted access to […].(cited from Markert and Nissim 2006: 159, British National Corpus)

       b.   At the time of <u>Vietnam</u>, increased spending led to inflation and a trade deficit. (*ibid.* p. 158)

       c.   We drank a nice <u>Bordeaux</u> last night.

The associations underlying such metonymies are *conventionalized*, in the sense that they are either memorized as lexical items or the lexical process that generates them is fast and effective. In this article, we consider them instances of 'regular metonymy' because they reflect recurrent, entrenched conceptual mappings such as PART FOR WHOLE, CAUSE FOR EFFECT, PERSON FOR ROLE, PLACE FOR EVENT, etc. However, other metonymies, such as the ones illustrated in (1), are not widely conventionalized (Markert & Nissim 2006) although they use the same mental processes. Many of them are produced and comprehended online; that is, they are novel and their interpretation depends on the concrete context of the utterance together with linguistic and pragmatic principles of interpretation. These utterances can be deemed unacceptable if the context in which they are produced is not taken into account, such as (1a). Following Markert and Nissim (2003, 2006) we will label such utterances as 'unconventional metonymy,' being

mindful of the fact that regular and novel metonymy are not mutually exclusive, but rather two opposites on a cline of metonymy conventionalization.

What happens when acquiring this lexical process in a second language? Cognitive linguistics (Lakoff and Johnson 1980; Lakoff 1987, Kövecses and Radden 1998) claims that metonymy is a universally available mechanism, although it is dependent on conceptual structures, cultural knowledge, and pragmatic routines. Jackendoff's (1997) conceptual semantics framework and Pustejovsky's (1995) generative lexicon also consider this mechanism of semantic enrichment to be universally available. These three approaches to metonymy are quite different, yet they all predict that comprehending metonymic transfer of meaning would be easy for second language learners; once familiar with the lexical items used in the metonymic expressions, they should be able to compute the additional meaning through a cognitive process they transfer from their native grammar. However, metonymy comprehension is a non-compositional meaning computation, so we would expect to see some higher processing costs incurred in comparison with compositional meaning computations. We could expect to see an effect of regularity, in the sense that learners may be more familiar with the metonymic patterns that occur more consistently and predictably. Processing costs will be a factor in the regularity effect: The more regular and predictable the pattern of metonymy, the easier it will be to process. We also expect there may be different influences from different L1s as some languages may employ regular metonymy but refrain from novel, productive metonymy. These are all empirical questions that we address in this experimental study.

In an effort to stay away from culturally-specific knowledge (e.g., not all L2 learners of English may know that Raisin Bran is a popular brand of breakfast cereal), we investigated only metonymic mappings containing constituents that are part of current global culture. We

experimented with research designs and test items, piloting widely among English native speakers before translating our tests into Spanish and Korean. An 'easier' paraphrase task and a 'harder' acceptability-in-context judgment task were finally selected for the experimental study. Regular meaning shifts were designed to compare with more productive sense shifts. We compared acceptable metonymic patterns with unacceptable ones and with baseline sentences without metonymies. Our main findings reveal that there is a lot more speaker variation than the literature reviewed here acknowledges or is aware of, and that regular metonymic shifts are rated as much more acceptable than novel, productive shifts, but only in English. While regular metonymy patterns are not as conventionalized in Spanish and Korean as they are in English (see Cuenca & Hilftery 1999 on Spanish metonymy; Park 2011 on Korean metonymy), the L2 English learners in the present study are eventually capable of achieving and even surpassing English native speaker patterns of behavior in terms of interpreting and accepting or rejecting regular and novel metonymy. Teachers of second language should be aware of these findings because they support existing research (e.g. Hashemian & Nezhad 2007) which indicates that engaging students in metonymy recognition and explicit discussion at all levels of proficiency may promote faster and more accurate understanding of figurative language.

THEORETICAL ACCOUNTS AND EXPERIMENTAL EVIDENCE OF METONYMY COMPUTATION

In order to provide a broad overview of the different approaches to metonymic analysis, the predictions that these approaches yield for L2 metonymy acquisition, and the small body of metonymy processing and acquisition data, we start by outlining the way metonymy has been

analyzed by cognitive linguistics, by the radical pragmatics approach, and by generative linguists.

Lakoff and Johnson (1980) proposed that metaphor and metonymy should be considered as the fundamental mechanisms of language and thought and subsequently, in the last two decades, metonymy has been studied widely in cognitive linguistics (Langacker 1993; Kövecses and Radden 1998; Radden and Kövecses 1999; Panther and Thornburg 2003; Panther, Thornburg, and Barcelona 2009; Benczes, Barcelona, and Ruiz de Mendoza 2011). We cannot do the vast literature on metonymy within cognitive linguistics justice, so we limit our discussion here to the approach's basic underpinnings. The major features of the cognitive linguistics analysis are the fundamentally conceptual nature of metonymy, the fact that it is grounded in experience, and that it involves experientially and conceptually connected 'contiguous' elements (Barcelona 2011: 48-49). A useful way for describing various cases of metonymy within this theoretical approach is the continuum of metonymicity, based on the display of more or less prototypical metonymy properties. Since this approach to metonymy is grounded in the description of cognitive shifts and thought processes, it has not produced much experimental data on how speakers of a language process the various types of metonymy. According to the cognitive approach, it is possible that various levels of metonymicity could be displayed in different languages.

One of the most important questions related to the production and comprehension of metonymy is what makes some metonymies viable, so attested in human language, while others are not. In other words, descriptions of cognitive rules and processes should not overgenerate, that is, allow unattested, infelicitous metonymies into the grammar. Various approaches treat this issue differently. If the notion of contiguity is at the core of most definitions of metonymy,

traditional grammar approaches locate contiguity relations in the real world while cognitive approaches locate it at the conceptual level (Radden and Kövecses 1999). Lakoff (1987) accounts for metonymic contiguity within the framework of Idealized Conceptual Models (ICM). Accepting that the latter are the best description of metonymy because they include not only people's encyclopedic knowledge but their cultural knowledge as well, Radden and Kövecses (1999) explore which ICMs can give rise to metonymic transfer. They propose that stereotypical associations are conceptualized, while non-stereotypical ones are not (p. 22).

Panther and Thornburg (2003), still within a cognitive linguistics approach, propose that the comprehension and generation of metonymy is guided by certain 'inferential principles' or 'conceptual metonymies' such as PART-WHOLE, CAUSE-EFFECT, PERSON-ROLE, and REPRESENTATION-REPRESENTED. They consider these reasoning principles to be conceptual devices not restricted to language, but also used in other semiotic systems. If a substitution is attempted that is outside of these ready-made and accessible inferential principles, it will not be comprehended adequately. The prediction of this kind of approach for L2 acquisition is that the reasoning principles should be universal, especially because they are not restricted to (a specific) language, hence available to transfer or utilize directly.

The radical pragmatic approach to metonymy, essentially a Relevance Theory approach, (Nunberg 1979, 1995; Papafragou 1996) has been among the first to tackle the question of why some shifts of meaning are attested and comprehensible while others do not make sense. According to this approach, a change of sense is allowed only when there is an important relation between the two meanings, based on the interlocutors' knowledge of the world and a particular context (Nunberg 1995). For example, 'Put Shakespeare on the top shelf' is acceptable because there is a noteworthy link between the work and the author, within the PRODUCER FOR PRODUCT

regular polysemy. Papafragou (1996) further singles out salience in the discourse situation as the important relationship supporting metonymy construal. In addition, Nunberg (1995) proposes that there are language-specific polysemy conventions, which would presumably make the metonymies sanctioned by them easier to process than more irregular, rarer uses. Thus the radical pragmatic approach would predict that some restructuring of the grammar would be necessary in L2 acquisition, coupled with 'noticing' that discourse salience supports novel metonymy.

Generative linguistics views sense shifts within the polysemy of a lexical item as separate lexical rules. Jackendoff (1997, 2002), for example, integrates inferential rules into the Conceptual Structure component of his language architecture. Following Pustejovsky (1995), he analyzes the 'ham sandwich' metonymy as a case of 'enriched composition', where the ham sandwich stands for 'the person *contextually associated with* a ham sandwich'. Other comparable lexical processes include complement coercion ('The author began the book in 1997' meaning 'the author began *writing* the book') and aspectual coercion ('The lighthouse flashed until dawn' where the punctual predicate 'flash' is coerced by the time interval adverbial to mean a series of punctual events). What unites these enriched composition cases is the presence of silent meaning that is not read off of syntactic structure, but is ostensibly supplied by the context. That is, compositionality is evidently violated. Importantly for us in this article, this approach considers conventionalized and language-specific aspects of the general processes of meaning transfer to be expressed and learned as lexical rules. These rules serve as language-specific filters on the general, universally available cognitive process. The prediction of this type of treatment is that even if the universal process of reference transfer is available, the language-specific rules may take time to acquire in a second language.

Experimental evidence on the processing of metonymy is most relevant to evaluating the various treatments of metonymy in the theoretical literature. The first attempts at comparing metonymic expressions with baseline literal expressions using the eye-tracking technique (Frisson and Pickering 1999) did not find that PLACE FOR EVENT and PLACE FOR INSTITUTION metonymic shifts were hard to process for native English speakers. The authors argued that the processor can obtain metonymic interpretations without appreciable difficulty. However, these two metonymic shifts are what we call regular metonymy, as opposed to novel metonymy, and their interpretation is not context-dependent. Recently, Rundblad and Annaz (2010) studied the development of metonymy comprehension compared with metaphor comprehension in English speakers aged 5;3 to 37;1, using an off-line interpretation task. They discovered that metonymy comprehension develops over time, reaches ceiling around the age of 12, and is significantly correlated with age and vocabulary development. More in line with our research questions in this article, Frisson and Pickering (2007) tracked eye movements to study the effect of familiarity on the processing of the PRODUCER FOR PRODUCT regular metonymic pattern. They found that familiar metonyms (e.g., 'read Dickens') were straightforward to process, but unfamiliar metonyms (e.g. 'read Needham') caused processing difficulty unless context made it clear that the metonymic interpretation would be appropriate.[1]

Finally, let us review what we know about the L2 acquisition of metonymy. This research is almost exclusively done within the cognitive linguistics framework. Littlemore (2009) and Barcelona (2010) highlight the importance of understanding figurative language, including metaphor and metonymy, for full communicative competence to be acquired by L2 learners. However, empirical work for the purpose of establishing what learners comprehend and produce is very much needed. Among the notable recent studies are Chen and Lai (2011), which

looks at Chinese native speakers' awareness of figurative language in L2 English, as well as Littlemore, Trautman Chen, Koester, and Barnden (2011), a study of non-native speakers' English metaphor comprehension. The participants in the latter study were from various native language backgrounds, and were attending a British university as undergraduates. They identified words or phrases that they considered difficult in a sample of academic texts. The authors analyzed the sources of difficulty and found that, among the expressions composed of words known to the participants, around 40% were metaphoric expressions. Since metaphor computation is cognitively close to novel metonymy, this finding suggests that non-compositional meaning enrichment increases the cognitive load of L2 speakers and leads to (possibly undetected) errors of understanding.

To summarize our necessarily brief review of several theoretical approaches and areas of investigation, metonymy is viewed as a productive lexical mechanism whereby one (literal) sense of a word is extended to another, related and contextually supported, sense. This cognitive process is experientially based and relies on a universal meaning computation mechanism, but may differ in regularity, scope, and frequency of usage across languages of the world. It is important to distinguish between L2 speakers' comprehension of unconventional and more regular metonymic patterns because (English) native speakers process them differently, and because the patterns produce different cognitive loads.

RESEARCH QUESTIONS AND PREDICTIONS

We conceive of our research questions within the larger theoretical import of which properties pose difficulty to acquire and which properties come for free in L2 acquisition. Basing her arguments on Jackendoff's (2002) view of linguistic differences, Slabakova (2008) formulated

the Bottleneck Hypothesis (BH), arguing that functional morphology poses the highest difficulty in L2 acquisition, because it reflects all the crucial morpho-syntactic features in a language. An integral part of the BH is the claim that universal meanings will not pose insurmountable difficulty to L2 learners because they will come for free from the native language. However, universal meanings or cognitive patterns can incur more cognitive effort to acquire if their expression is different in the L2, if they are rare in the input to learners, or if they are computationally complex.

With respect to regular versus unconventional metonymy more generally speaking, we contend the former will be acquired as a well-supported, frequent lexical rule, while the latter will be acquired as a productive but less regular and rare lexical rule creating new associations. Furthermore, we are aware of the fact that our division of metonymic patterns into regular, less regular, and unacceptable is based on English, the target language. Thus the patterns that are regular in English may be less regular in Spanish and Korean. In the absence of corpus data for the metonymic patterns investigated in this experiment, we rely on native speaker judgments to establish regularity through acceptability ratings.

Being mindful of the considerations above, the findings in the literature, and using Slabakova's (2008) BH logic, we predict the following for L2 regular and unconventional metonymy:

With respect to regular metonymy:

1. If the learners' native languages attest a metonymic pattern as regular, learners will judge that pattern in English correctly even at intermediate proficiency levels.

2. If, however, there is a L1-L2 difference in regularity, as indicated by acceptability ratings, we expect acquisition to be harder, but still ultimately successful. The reason for the

latter prediction is that regular metonymic pattern will be freely available in the input to L2 learners.

With respect to less regular, unconventional metonymy:

3. If the learners' native languages attest an unconventional metonymic pattern to the same extent as English does, learners will judge that English pattern correctly, even at intermediate proficiency levels. However, acceptability ratings may differ among individuals, reflecting the rarity and unconventionality of the pattern.

4. If a less regular metonymic pattern is attested in English but is more acceptable or not as acceptable in Korean/Spanish, in other words, if there is a mismatch of acceptability for a rare pattern:

a. Intermediate proficiency learners will struggle with those metonymic patterns; their judgments will differ from those of native speakers.

b. Advanced learners will be able to overcome transfer from the native language for both regular and novel metonymy.

These predictions reflect the interplay between regularity of a metonymic pattern with its frequency in the input to learners and the potential mismatch between L1 and L2 in terms of pattern availability.

THE PRESENT STUDY

Participants

One hundred and forty-six individuals participated in our experimental study. Some of the participants were compensated with extra credit or monetarily for their time, others volunteered to take the tests. They were divided into three control groups (English, Spanish and Korean) and

three learner groups, the latter based on their proficiency in English. See Table 1 for a summary

of the participant information. An excerpt from the Oxford Test of English was modified to

establish proficiency levels among the learners. The test included 50 grammar-based multiple-

choice test sentences, with a maximum score of 50. The mean scores and ranges are displayed in

Table 1.

[TABLE 1 NEAR HERE]

Comparing the means and ranges of the two advanced groups, it can safely be assumed

that they are not equal in proficiency in English, with the Spanish advanced group more

proficient than the Korean advanced group. In essence, we have three groups differing in

proficiency, arranged as follows: Korean intermediate learners < Korean advanced learners <

Spanish advanced learners.

Selection of novel and regular metonymic patterns

It is important to address the issue of the relative degree of conventionality of the metonymic

patterns, and more specifically, from whose point of view this attribution is made.[2] Admittedly,

we took an expedient approach to this decision. The literature on metonymy processing (Frisson

and Pickering 1999, 2007) has documented the possible differentiation of familiar and non-

familiar metonymic patterns in English, but no data exist with respect to this differentiation in

Spanish and Korean. Cross-linguistic work on metonymy patterns focuses more on inter-

language differences within one or another pattern, but not so much on measuring degrees of

conventionality within a language. For reasons of lack of data from Spanish and Korean, and in

the interest of a tightly focused research design, we divide the metonymic patterns into novel and

regular from the point of view of English. However, the ratings of the Korean and Spanish native

speakers will help us to establish a preliminary picture as to which pattern might be conventionalized in these languages.

Main tests and experimental materials

Two pen-and-paper tasks were compiled in English and then translated from English into idiomatic, natural Spanish and Korean by native speakers of those languages. Test sentences were taken mainly from the existing literature on metonymy. The decision to translate the test items was taken so that we could control the metonymic patterns we were testing.[3] Piloting with 5 native speakers allowed us to retain only the most (un)acceptable sentences for the final tests.

Also by design, the two tasks differed in difficulty so that a range of possible interpretations could be detected. We intended The Paraphrase Task (see example (3)) to be accessible to intermediate proficiency learners, since it asked participants to choose from four given interpretations. One paraphrase corresponded to the metonymic interpretation and one corresponded to the literal interpretation; participants also had the possibility to indicate that both interpretations were available, or that neither interpretation was acceptable. In a very small percentage of the test items, (c) was the correct answer, as both the literal and the metonymic paraphrase were acceptable.

(3)   The first violin has the flu.

    a.   The first violinist is sick with the flu.   ← expected

    b.   The first violin is not working properly because of an illness.

    c.   Both meanings are possible.

    d.   Neither meaning is possible.

In this task, we tested four metonymic patterns. Two of those we considered rarer

patterns based on our own piloting: INSTRUMENT FOR AGENT (n=4) as in (3) above, and another, which we termed LOOSE ASSOCIATION (n=9), comprising saliently but temporarily related events and people (e.g., 'The hamburger is a lousy tipper'). The two other patterns were considered regular and more conventionalized: POSSESSOR FOR POSSESSED (n=4) (e.g., 'Did you notice that you have a flat tire?') and PRODUCER FOR PRODUCT (n=4) (e.g., 'Proust is on the top shelf.') The test also included the same number of baseline sentences without metonymy (n=21)[4] and 14 fillers.

The second task was an Acceptability Judgment Task, in which participants rated a sentence in the context of another sentence using a scale of 1 to 5 as in (4), where 1 was 'unacceptable' and 5 was 'perfectly acceptable'.

(4)   The soldiers began to move up the field during battle. The colonel gave the cannon the signal to fire.

| 1 | 2 | 3 | 4 | 5 |

In this task, we included four types of test items: baseline (without metonymy), experimental (good), experimental (bad), and fillers. The acceptable experimental sentences were divided into three metonymic patterns: INSTRUMENT FOR AGENT, LOOSE ASSOCIATION, and PRODUCER FOR PRODUCT (n=4 in each). These sentences were paired with others containing metonymies that violated metonymic requirements, see example (5). In creating these test items, we were led by the findings of Rabagliati, Marcus, and Pilkkänen (2011). Using paraphrase judgments, these researchers investigated whether similarity, noteworthiness, cue validity (frequency of association), or salience were good predictors for the acceptability of a shifted sense. The only factor that showed some predictive effect on acceptability, though not very robust, was noteworthiness. Nunberg defines a property as noteworthy "if it offers a useful way

of classifying its bearer relative to the immediate conversational interests" (Nunberg 1995: 114). We constructed our unacceptable test items such as (5) with the intention that they violate this factor.

(5)   The small boy threw a hammer in the garage while he was playing. The workbench suffered a mental breakdown.

                    1   2   3   4   5

In this example, there is a meaning shift, in the sense that 'the workbench' is a thing while suffering a mental breakdown is a characteristic of people. However, 'the workbench' in (5) cannot be usefully shifted to a person in classifying the conversational situation, hence it is not noteworthy.[5] These test sentences were included to probe deeper into our predictions on the productivity of unconventional metonymy, since knowledge of a productive pattern includes rejecting sentences that do not obey the pattern. In addition to the acceptable and unacceptable metonymy items, there were 12 acceptable baseline sentences without metonymy and 14 fillers. If this research design succeeded in capturing speakers' intuitions, we expected to see statistically significant differences, at a minimum, between baseline items and acceptable metonymies, on the one hand, and unacceptable metonymies, on the other. Reliability statistics on the two tasks obtained a Cronbach Alpha value of .738 for the Paraphrase Task and .853 on the Acceptability Judgment Task, which are statistically significant positive correlations.

RESULTS

Paraphrase Task

Recall that in the Paraphrase Task participants had to choose between four options, so chance performance is 25%. In the figures that follow, we display accuracy results (out of one). When

the speaker chose the correct paraphrase, one point was given; in all other cases, no point was given. The medians and ranges of the English native speakers and the learners on the two regular metonymy patterns and the baseline (test sentences without metonymic interpretation) are plotted as clustered boxplots in Figure 1. The size of the boxes indicates the range from the 25th to the 75th quartile, or the interquartile range (IQR). The whiskers represent the lowest datum within 1.5 IQR of the lower quartile, and the highest datum within 1.5 IQR of the upper quartile. Asterisks in the graphs represent extreme outliers while circles represent mild outliers. It is quite clear from the chart that there is at least some non-normality in the data in terms of distribution, because some outliers are shown in all the groups and some of the whiskers for the Korean groups span the whole range. Appreciably, the advanced Spanish learners of English are as accurate as the native speakers on these conditions.

[FIGURE 1 NEAR HERE]

In Figure 2, we have plotted the same baseline accuracy for the four groups contrasted with the two conditions that we stipulated would be less conventionalized and more ad-hoc: INSTRUMENT FOR AGENT and LOOSE ASSOCIATION. The ranges especially of the INSTRUMENT FOR AGENT condition (in blue) are much larger than the baseline. Furthermore, the advanced Spanish group presents a marked parallel with the native speakers, displaying even less variation than the latter.

[FIGURE 2 NEAR HERE]

We have not represented the ratings of the Spanish and Korean native speakers in these graphs, because results become too difficult to compare visually. However, we include these groups in the statistical analyses because we are interested in the effects of the native language on L2 comprehension. A Repeated Measures ANOVA was conducted on the mean correct

17

choices of all conditions and all groups in the Paraphrase Task. There was a significant effect of Condition ($F_{4, 137} = 12.921$, $p < .0001$), a significant effect of Group ($F_{5, 140} = 23.306$, $p < .0001$), and a significant Condition * Group interaction ($F_{20, 560} = 3.055$, $p < .0001$), the latter suggesting that not all groups are performing equally on all conditions.

We repeated the ANOVA, separating what we considered the regular metonymy (POSSESSOR, PRODUCER) from the novel metonymy (INSTRUMENT, LOOSE ASSOCIATION) conditions, and compared them with the baseline. When the two regular metonymy conditions combined were compared with the baseline, there was no effect of Condition ($F_{1,140} = .126$, $p = .723$), a significant effect of Group ($F_{5,140} = .29.849$, $p < .0001$), but no Condition*Group interaction ($F_{5,140} = 2.049$, $p = .075$). When the two novel conditions were compared with the baseline, there was an effect of Condition ($F_{1,140} = 81.426$, $p < .0001$), an effect of Group ($F_{5,140} = 21.893$, $p < .0001$), and a Condition*Group interaction ($F_{5,140} = 7.088$, $p < .0001$). Thus we can safely assume that the participant groups treat ad-hoc, novel metonymy differently from regular metonymy, as compared with baseline sentences without metonymy in the Paraphrase Task.

We probed further to see which individual groups comprehended regular and novel metonymy differently with respect to the baseline sentences. For this calculation, we combined the two regular metonymy and the two novel metonymy accuracy means for each participant group (out of two) and compared them with the baseline accuracy. The paired samples *t*-tests are presented in Table 2.

[TABLE 2 NEAR HERE]

The comparisons in Table 2 are illuminating. First of all, the table shows that English and Korean native speakers choose both novel and regular metonymy interpretations accurately, and do not distinguish them from baseline acceptable sentences without metonymy. However, that is

not the case for the Spanish native group, which does not find novel metonymy as acceptable as the baseline. Next, Spanish advanced learners have clearly overcome the native bias in their L2, because their behavior is indistinguishable from that of the English native speakers on the novel conditions. Finally, although the Korean native speakers choose novel metonymy freely in Korean, this is not true for the Korean learners of English. Both Korean learner groups demonstrate that novel metonymy is more difficult for them because they choose it less accurately than baseline, while they treat regular metonymy and baseline the same way. In other words, the Korean learners have not reached native-like levels in their L2 English, while the Spanish learners of English have already achieved target-like L2 competence.

Acceptability Judgment Task

Recall that the Acceptability Judgment Task was considered more difficult because the participants had to compute an interpretation in context and evaluate the acceptability of the whole utterance in that context. Furthermore, they had a scale of one to five to rate their acceptability, which allowed them to provide more nuanced judgments than on the Paraphrase Task. Not surprisingly, the findings of this task provide us with a more detailed picture of how acceptable participants find the various types of metonymy.

Within the test sentences with metonymy, three patterns were included: INSTRUMENT FOR AGENT, LOOSE ASSOCIATION, and PRODUCER FOR PRODUCT. It was hypothesized that the baseline sentences would be rated high and the unacceptable test sentences low, on a scale of 1 to 5. Ideally, the sentences with metonymy would be rated as highly as the baseline sentences. In the next four figures, we present boxplots of each acceptable metonymy condition and the unacceptable metonymy test sentences.

[FIGURE 3 NEAR HERE]
[FIGURE 4 NEAR HERE]
[FIGURE 5 NEAR HERE]
[FIGURE 6 NEAR HERE]

We highlight several points in Figures 3 to 6. First of all, the acceptability judgments of the three native speaker groups involve great variability, which suggests that the judgments are not stable and consistent across speakers. Second, one of the two novel metonymy patterns, INSTRUMENT FOR AGENT, has acceptability ratings in the lower to middle parts of the scale (see Fig. 3), meaning that speakers do not find these sentences highly acceptable, even in context. On the other hand, the LOOSE ASSOCIATION pattern (Figure 4) is more acceptable. Third, there is a marked contrast in this respect between the ratings on the novel and the regular metonymy patterns, at least for the English and Korean native speakers. Table 3 provides the means and standard deviations on all experimental patterns.

[TABLE 3 NEAR HERE]

It is not surprising that English learners do not rate unconventional metonymy highly, as Table 3 illustrates, given that the English native speakers do not rate it highly. Spanish advanced speakers are the only learner group that patterns with the English native speakers in all respects. The Korean groups' behavior differs both from the Korean natives' and the English natives' behavior, which is perhaps explained by their lower proficiency level. We come back to these findings in the discussion.

An ANOVA (RM) on all the conditions (all metonymic patterns including the unacceptable one) in this task reveals a significant effect of Condition ($F_{4,136} = 124.99$, $p < .0001$), an effect of Group ($F_{5, 139} = 2.350$, $p = .044$), and a significant Condition*Group interaction ($F_{20, 556} = 10.458$, $p < .0001$). These results indicate that the six groups do not treat all the metonymy and baseline conditions equally. One-way ANOVAs show a main effect of

20

Group in all acceptable metonymy conditions but the LOOSE ASSOCIATION condition (in the

INSTRUMENT FOR AGENT condition $F_{5,144} = 4.190$, $p = .001$, in the LOOSE ASSOCIATION condition

$F_{5,144} = 1.140$, $p = .342$, in the PRODUCER FOR PRODUCT condition $F_{5, 144} = 13.067$, $p = .0001$).

Perhaps more interesting from the point of view of our research questions, we would like

to establish not only how the groups differ from each other but also what the variation is within

groups. In order to uncover whether each individual metonymy differs from the baseline and the

unacceptable test sentences within each language, we performed a series of paired samples *t*-

tests. The results are shown in Table 4 in Appendix A.

[TABLE 4 NEAR HERE]

Ideally, we would like all experimental metonymy patterns to be rated the same as the

baseline condition but differently from the unacceptable one. However, that is not what we find.

While—as Table 4 reveals—all native groups rate the purportedly acceptable metonymies

significantly higher than the unacceptable metonymy, thus validating our test instrument,

comparisons between the acceptable metonymies and the baseline bring to light some

unexpected outcomes. The English group rates only the PRODUCER FOR PRODUCT metonymy as

highly as the baseline test sentences. The Korean native group rates both PRODUCER FOR

PRODUCT and LOOSE ASSOCIATION as highly as the baseline. Finally, the Spanish native speakers

situate the ratings of all metonymy conditions somewhere between the acceptable baseline and

the unacceptable metonymy, with significant differences everywhere. We discuss the

implications of these findings in the next section.

DISCUSSION

English native speaker variation

The first impression on looking at the boxplots in Figures 1 through 6 is that there is a lot of variation in the metonymy data of the native speakers, more than is common when native speakers are experimentally tested on syntactic or semantic properties. Standard deviations on both tests and whiskers in the boxplots signifying ranges of scores are quite dramatic. While some individuals are quite tolerant of metonymic patterns, other individuals judge them more conservatively as relatively unacceptable.

Furthermore, it is not only the large interpersonal variation that is striking. The Acceptability Judgment Task allows us to evaluate the tolerability of novel metonymy directly on a scale of one to five. Although the vast literature on metonymy considers the process productive, our English native speakers (n=36) did not rate novel metonymy very highly (Table 3), even though they still distinguished it from unacceptable metonymy (Table 4). These findings suggest that shifting a meaning, even for native speakers and even if the metonymy is productive, incurs higher processing costs, as attested by individual variation and lower acceptability. This fact is important because we can expect that the higher processing costs will be reflected in (even more) variability and inaccuracy among the L2 speakers, independent of which approach to metonymy one assumes.

Variation among the English, Spanish, and Korean native speakers

The boxplots and Table 3 illustrate that the native speaker groups differed among each other. It seems that Korean is quite tolerant of novel metonymy, and maybe more tolerant than English with the LOOSE ASSOCIATION pattern. On the other hand, Korean speakers judge regular metonymy with a mean rating that is a full point lower than that of the English native speakers, suggesting that the pattern PRODUCT FOR PRODUCER may not be as conventionalized in Korean as

it is in English. However, we must also be mindful of the fact that Korean native speakers were very accurate on the regular metonymy in the Paraphrase task, such that the results of the two tests seem to contradict each other. This apparent contradiction can be resolved if we recall that the Paraphrase Task offered ready-made interpretations to choose from, while the harder Acceptability Judgment Task asked speakers to evaluate metonymy. If regular metonymy in Korean is not highly conventionalized but is still productive, its relatively lower acceptability in the harder task may be plausibly explained with its higher demand on processing resources. This is an empirical question that will be addressed in future research.

Table 2 shows that Spanish native speakers are somewhat less tolerant to novel metonymy patterns than English native speakers, while they accept regular metonymy, in the Paraphrase Task. Figures 3 and 4 illustrate that Spanish native speakers pattern similarly to the native English speakers on the novel metonymies in the Acceptability Judgment Task, but the surprising finding is their relatively low rating of regular metonymy as seen in Figure 5 and Table 3. Performance on the harder task suggests that, as in Korean, regular metonymic patterns in Spanish may not be as highly conventionalized as they are in English. In sum, there is a match between English, Spanish, and Korean with respect to the reduced acceptability of the unconventional patterns but a mismatch with respect to the regular pattern. We discuss these results in more detail in Slabakova, Cabrelli Amaro and Kang (2013).

Regular versus less regular patterns among native speakers?

Another important question that our experiment sought to address is whether regular and novel metonymic patterns were treated differently by the native speakers of Spanish, Korean, and English. The literature on metonymy processing (Frisson and Pickering 1999, 2007) has

documented this possible differentiation of familiar and non-familiar metonymic patterns in English. In the easier Paraphrase Task, the native participants showed no effect of Condition with the regular metonymy but a significant effect of Condition for the novel metonymy. However, looking at the results on the harder Acceptability Judgment Task, the emerging picture is different. The within-group comparisons in Table 4, Appendix A address this question best. For the English native speakers, only the PRODUCT FOR PRODUCER metonymy is as acceptable as the baseline. Koreans evaluate PRODUCT FOR PRODUCER but also LOOSE ASSOCIATION as equally acceptable as the baseline, while the Spanish native speakers rate none of the metonymic pattern as highly as the baseline. We can conclude that only the English native speakers treat novel and regular metonymy differently, with the latter highly conventionalized, while the other native groups do not (for more details see Slabakova *et al*. 2013).

Learner behavior

We are now in a position to come back to our research questions regarding L2 learner behavior. Recall our prediction with respect to the regular patterns depended on whether or not we would uncover high regularity in Spanish and Korean. Since we did not, prediction 2 of successful acquisition is expected to obtain, although with somewhat lower evaluations. This is exactly what our experimental findings attest: While the numerical evaluation ratings are somewhat depressed (see Table 3), all of the learner groups treat regular metonymy and baseline sentences similarly. As suggested above, at least one reason for this finding may be the availability of supporting evidence for this pattern in the input learners encounter in English. Whether this is truly the case is an empirical question that will be tested going forward by controlling for attested frequency of particular patterns in the input.

Our next two predictions depended on the acceptability of the novel metonymic patterns in English, Korean, and Spanish. All these languages accept novel metonymy, but not to the same extent. Spanish and English native speakers evaluate novel metonymy as less acceptable than the baseline, but Korean native speakers deem one of the novel metonymy patterns and baseline sentences as equally acceptable (Table 4). Prediction 3 describes the situation we found for the INSTRUMENT FOR AGENT unconventional metonymic pattern: None of the native groups are particularly accepting of this pattern. It is hardly surprising, then, that the learners have very similar behavior (compare Table 3 and Table 4).

Prediction 4 describes a case where a less regular metonymic pattern is more acceptable than in English. We do have such a situation in our data as well: the case of Korean and English native speakers on the LOOSE ASSOCIATION pattern. As Table 4 indicates, Korean natives do not distinguish this pattern from baseline sentences while English native speakers do. Both the intermediate and the advanced Korean learners transfer the same predilection for this less regular metonymic pattern (Spanish learners have nothing to transfer, as their native language and the target language are matched in this respect). Thus, Korean learners are less than fully target-like with this pattern, because they have to preempt a native pattern, or learn its relative unacceptability, for which they might not have enough evidence in the input. In any event, preempting a native construction is considered eminently difficult in L2 acquisition (e.g., White 1991, Trahey and White 1993). However, we should also keep in mind that even the advanced Korean-native learners of English are not as advanced as the Spanish group. Thus it is possible in principle that acquiring the relatively low acceptability of the LOOSE ASSOCIATION pattern will come in the future for those learners. There is a significant body of work on the L2 acquisition of meaning (Slabakova 2006), that has documented overcoming of learnability problems, so

successful preemption should be feasible in principle.

Our general prediction about possible successful acquisition is clearly supported by the behavior of the Spanish advanced learners. It is true that these learners are highly proficient in English, but they have had to overcome severe restrictions on metonymy comprehension coming from their native Spanish. This is most clearly visible in Table 3, where the Spanish native evaluations of all metonymic patterns fall between the baseline and unacceptable metonymy, while the advanced learners of English exhibit exactly the native English pattern of evaluations, with a significant distinction between novel and regular metonymy. All in all, theirs is a complete success story.

Our predictions were partly based on the logic of the Bottleneck Hypothesis (Slabakova 2008) within the generative SLA framework and we should discuss how these predictions are supported by the results on the whole. Recall that the Bottleneck Hypothesis contends that universal semantic and pragmatic properties will come for free in the grammar of learners, transferred from their native language. There is clearly evidence for L1 transfer in our data, insofar as the non-proficient learners' tolerance of metonymy goes only as far as their native language tolerance. At later stages of development, learners are struggling with patterns not sufficiently supported by the input, hence incurring higher processing costs because they feel unfamiliar and unrehearsed. The superior performance of the very advanced learners is also consistent with the Bottleneck Hypothesis logic, because the transferred mechanism of non-compositional computation, coupled with observation of the L2 target input, inevitably leads to successful acquisition.

Implications of the findings

What do these patterns of behavior tell us about the L2 acquisition of universal lexical processes? First, the lexical computation that we investigated in this study, metonymic shift of meaning, did not turn out to be so universal in terms of high acceptability and accurate recognition, which could be indicative of *some* productivity but *lower frequency* for novel metonymy in English and Spanish. We have to conclude that even though a lexical process is attested in a language, what is of utmost importance for its L2 acquisition is how often it is used, to what classes of words it is applied, and which patterns of metonymy are conventionalized. Since English native speakers did not rate novel metonymy very highly, that type of metonymy is probably not very well supported with useful form-meaning mappings in the input available to learners of English. We have to appreciate that every metonymy has a literal and a shifted interpretation, which makes it cognitively demanding because its understanding necessitates accessing two related meanings, evaluating them based on context, and choosing one over the other as most fitting of the situation. With such learning tasks, exposing learners to ample unambiguous context and explicit instruction may be beneficial for reinforcing the productive mental operation that they can transfer from their native language. While metonymy comprehension is a cognitively demanding process, it is demanding for native speakers and learners alike, and many advanced learners demonstrate they can master the process. The task of explicit instruction would be to bring up the intermediate learners faster on the road to full figurative language comprehension.

CONCLUSIONS

In this article, we presented results of two off-line comprehension tasks investigating the acceptability of novel and regular metonymy by native speakers of English, Korean, and

Spanish, as well as learners of English as a second language. We were interested in uncovering conventional-unconventional metonymy computation discrepancies, and whether these types of metonymy were treated differently in the different languages. The distinction between conventional and unconventional metonymy is discussed at length by the existing theoretical treatments of metonymy. We obtained further experimental support for this distinction, but only for English. Significant findings were the acceptability of novel metonymy in Korean and the lack of conventionalization effect for regular metonymy in Korean and Spanish. The results suggest that the L2 acquisition pathway of this mental process is not unusual: It starts with L1 transfer, goes through a period of grammar restructuring, and ends with potentially full acquisition. Because this is a cognitively demanding process, its universal productivity is complicated by higher processing demand. Various acceptability levels of the metonymy patterns are also a factor. For future studies of metonymy comprehension, we would like to highlight the importance of cross-linguistic examination of metonymic patterns in native and second languages, not only in comprehension but in processing as well.

References:

Barcelona, A. 2010. 'Metonymic inferencing and second language acquisition.' *AILA Review* 23: 134–154.

Barcelona, A. 2011. 'Reviewing the properties and prototype structure of metonymy' in Benczes, R., Barcelona, A., Ruiz de Mendoza Ibáñez, F. J. (eds.): *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View* (pp. 7-58). Amsterdam: John Benjamins, Amsterdam.

Benczes, R., A. Barcelona, and **F. J**. **Ruiz de Mendoza Ibáñez** (eds.) 2011. *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View*. Amsterdam: John Benjamins.

Chen, Y. and **H. Lai.** 2011. 'EFL learners' awareness of metonymy-metaphor continuum in figurative expressions.' *Language Awareness* 21/3: 235-248.

Cuenca, M. J. and **J. Hilferty**. 1999. *Introducción a la lingüística cognitiva* [Introduction to Cognitive Linguistics]. Barcelona: Ariel.

Frisson, S. and **M. Pickering.** 1999. 'The processing of metonymy: Evidence from eye movements.' *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 25/6: 1366-1383.

Frisson, S. and **M. Pickering.** 2007. 'The processing of familiar and novel senses of a word: Why reading Dickens is easy but reading Needham can be hard.' *Language and Cognitive Processes* 22/4: 595-613.

Hashemian, M. and **M .R. T. Nezhad**. 2007. 'The development of conceptual fluency & metaphorical competence in L2 learners.' *Linguistik online* 30.

Jackendoff, R. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Jackendoff, R. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford, UK: Oxford University Press.

Kövecses, Z. and **G. Radden.** 1998. 'Metonymy: Developing a cognitive linguistic view.' *Cognitive Linguistics* 9: 37-79.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind.* Chicago: The University of Chicago Press.

Lakoff, G. and **M. Johnson.** 1980. *Metaphors We Live By*. Chicago: The University of Chicago Press.

Littlemore, J. 2009. *Applying Cognitive Linguistics to Second Language Learning and Teaching.* Basingstoke: Palgrave Macmillan.

Littlemore, J., P. Trautman Chen, A. Koester, and **J. Barnden.** (2011). 'Difficulties in metaphor comprehension faced by international students whose language is not English.' *Applied Linguistics* 32/4: 408–429.

Markert, K. and **M. Nissim.** 2006. 'Metonymic proper names: A corpus-based account' in A. Stefanowitsch and S. Gries (eds.): *Corpus-Based Approaches to Metaphor and Metonymy* (pp. 158-174). Berlin: Mouton de Gruyter.

Markert, K. and **M. Nissim.** 2003. 'Corpus-based metonymy analysis'. *Metaphor and Symbol*, 18/3: 175–188.

Nunberg, G. 1979. 'The non-uniqueness of semantic solutions: Polysemy.' *Linguistics and Philosophy* 3: 143-184.

Nunberg, G. 1995. 'Transfers of meaning'. *Journal of Semantics* 12: 143–184.

Panther, K.-U. and **L. Thornburg.** (eds.) 2003. *Metonymy and Pragmatic Inferencing*. Amsterdam: John Benjamins.

Panther, K.-U., L. Thornburg, and **A. Barcelona.** (eds.) 2009. *Metonymy and Metaphor in Grammar.* Amsterdam: John Benjamins.

Papafragou, A. 1996. 'On metonymy.' *Lingua* 99: 169-195.

Park, C. 2011. 'The role of metonymy in the interpretation of Korean multiple subject constructions.' *Language Sciences* 33: 206-228.

Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Rabagliati, H., G. Marcus, and **L. Pylkkänen.** 2011. 'Rules, radical pragmatics, and restrictions on regular polysemy.' *Journal of Semantics* 28/4: 485-512.

Radden and **Kövecses, 1999. '**Towards a theory of metonymy' in K-U Panther and G. Radden (eds.): *Metonymy in Language and Thought* (pp. 17-59). Amsterdam: John Benjamins.

Rundblad, G. and **D. Annaz.** 2010. 'Development of metaphor and metonymy comprehension: receptive vocabulary and conceptual knowledge.' *British Journal of Developmental Psychology* 28/3: 547-563.

Slabakova, R. 2008. *Meaning in the second language*. Berlin: Mouton de Gruyter.

Slabakova, R. 2006. 'Learnability in the second language acquisition of semantics: A bidirectional study of a semantic parameter.' *Second Language Research* 22: 498–523.

Slabakova, R., Cabrelli Amaro, J. and **Kang, S.K.** 2013. 'Regular and novel metonymy in native Korean, Spanish, and English: Experimental evidence for various acceptability.' *Metaphor and Symbol* 28/4: 275-293.

Trahey, M. and **L. White.** 1993. 'Positive evidence in second language acquisition: Some long effects.' *Studies in Second Language Acquisition* 15: 181–204.

White, L. 1991. 'Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom.' *Second Language Research* 7: 133–161.

Table 1: Summary of participant information

Group	Language Tested	N	Average proficiency score (range)
L1 English (control)	English	36	
L1 Spanish (control)	Spanish	23	
L1 Korean (control)	Korean	19	
L1 Korean/L2 English intermediate	English	16	24.7 (15-30)
L1 Korean/L2 English advanced	English	24	35.8 (31-47)
L1 Spanish/L2 English advanced	Spanish	28	43 (30-50)

Table 2: All groups' mean ratings (out of a maximum score of two) and multiple comparisons between novel metonymy, regular metonymy, and baseline in Parapharase Task, p-values (Tukey HSD test)

	Novel metonymy	Baseline	Contrast p value	Regular metonymy	Baseline	Contrast p value
English NSs	1.58	1.82	.26 n.s.	1.76	1.82	.394 n.s.
Korean NSs	1.73	1.80	.263 n.s.	1.87	1.80	.050**
Korean Int.	.84	1.20	.002*	1.02	1.20	.122 n.s.
Korean Adv.	1.10	1.50	.0001*	1.48	1.50	.722 n.s.
Spanish NSs	1.17	1.78	.002*	1.80	1.78	.569 n.s.
Spanish Adv.	1.72	1.78	.20 n.s.	1.88	1.78	.030**

* significant at p = .05 where novel metonymy means are lower than baseline
** significant at p = .05 where regular metonymy means are higher than baseline

Table 3: Means and standard deviations (in parentheses) in the Acceptability Judgment Task

Groups	INSTRUMENT FOR AGENT	LOOSE ASSOCIATION	PRODUCT FOR PRODUCER	Unacceptable metonymy	Baseline
English NSs	2.472 (.855)	3.5 (.910)	4.638 (.639)	1.861 (1.099)	4.50 (.507)
Korean NSs	3.263 (.805)	3.842 (.764)	3.631 (.683)	2.842 (.898)	3.842 (.602)
Korean Int.	2.75 (.683)	3.25 (.774)	3.312 (.704)	2.022 (.042)	3.50 (.730)
Korean Adv.	2.91 (.503)	3.33 (.816)	3.458 (.884)	3.00 (.589)	3.833 (.564)
Spanish NSs	2.391 (.782)	3.438 (.845)	3.869 (.868)	1.949 (.897)	4.261 (.619)
Spanish Adv.	2.509 (.807)	3.51 (.768)	4.378 (.682)	2.022 (.896)	4.16 (.443)

Appendix A

Table 4: Differences between acceptability in the metonymy experimental conditions, baseline,

and unacceptable metonymy for all language groups

Group	Contrast	*P* value
English NSs		
	INSTRUMENT FOR AGENT − Baseline	.0001
	INSTRUMENT FOR AGENT − Unacceptable	.0001
	LOOSE ASSOCIATION − Baseline	.0001
	LOOSE ASSOCIATION − Unacceptable	.0001
	PRODUCER FOR PRODUCT − Baseline	.8 n.s.
	PRODUCER FOR PRODUCT − Unacceptable	.0001
Korean NSs		
	INSTRUMENT FOR AGENT − Baseline	.001
	INSTRUMENT FOR AGENT − Unacceptable	.046
	LOOSE ASSOCIATION − Baseline	.502 n.s.
	LOOSE ASSOCIATION − Unacceptable	.0001
	PRODUCER FOR PRODUCT − Baseline	.172 n.s.
	PRODUCER FOR PRODUCT − Unacceptable	.0001
Spanish NSs		
	INSTRUMENT FOR AGENT − Baseline	.0001
	INSTRUMENT FOR AGENT − Unacceptable	.026
	LOOSE ASSOCIATION − Baseline	.0001
	LOOSE ASSOCIATION − Unacceptable	.0001
	PRODUCER FOR PRODUCT − Baseline	.023
	PRODUCER FOR PRODUCT − Unacceptable	.0001
Korean intermediate learners		
	INSTRUMENT FOR AGENT − Baseline	.0001
	INSTRUMENT FOR AGENT − Unacceptable	.0001
	LOOSE ASSOCIATION − Baseline	.52 n.s.
	LOOSE ASSOCIATION − Unacceptable	.0001
	PRODUCER FOR PRODUCT − Baseline	.248 n.s.
	PRODUCER FOR PRODUCT − Unacceptable	.0001
Korean advanced learners		
	INSTRUMENT FOR AGENT − Baseline	.0001
	INSTRUMENT FOR AGENT − Unacceptable	.681 n.s.
	LOOSE ASSOCIATION − Baseline	.002
	LOOSE ASSOCIATION − Unacceptable	.036
	PRODUCER FOR PRODUCT − Baseline	.036
	PRODUCER FOR PRODUCT − Unacceptable	.12 n.s.
Spanish advanced learners		
	INSTRUMENT FOR AGENT − Baseline	.0001
	INSTRUMENT FOR AGENT − Unacceptable	.0001

	LOOSE ASSOCIATION − Baseline	.0001
	LOOSE ASSOCIATION − Unacceptable	.0001
	PRODUCER FOR PRODUCT − Baseline	.025**
	PRODUCER FOR PRODUCT − Unacceptable	.0001

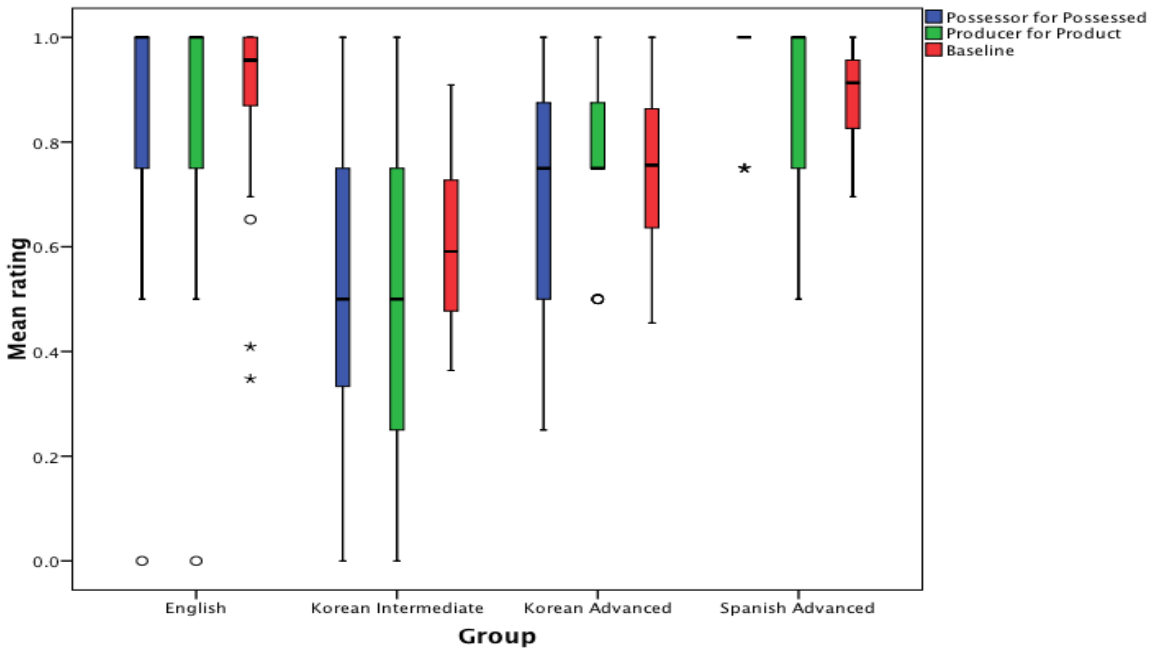** significant at $p = .05$ where regular metonymy means are higher than baseline

Figure 1: Clustered boxplots for the two regular metonymy conditions and the baseline in the
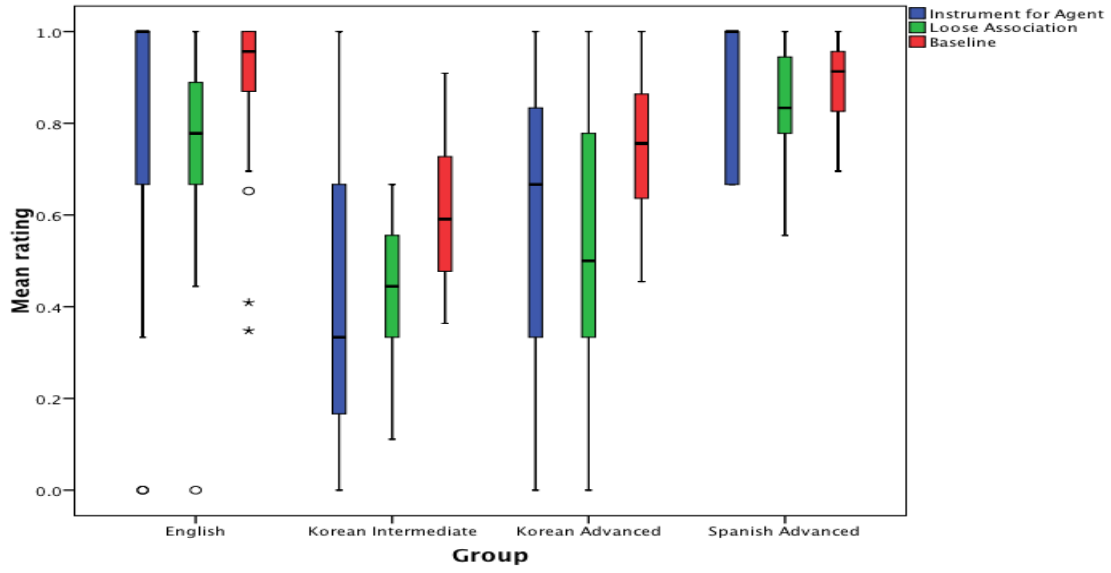
Paraphrase Task

Figure 2: Clustered boxplots for the two novel metonymy conditions and the baseline in the

Paraphrase Task
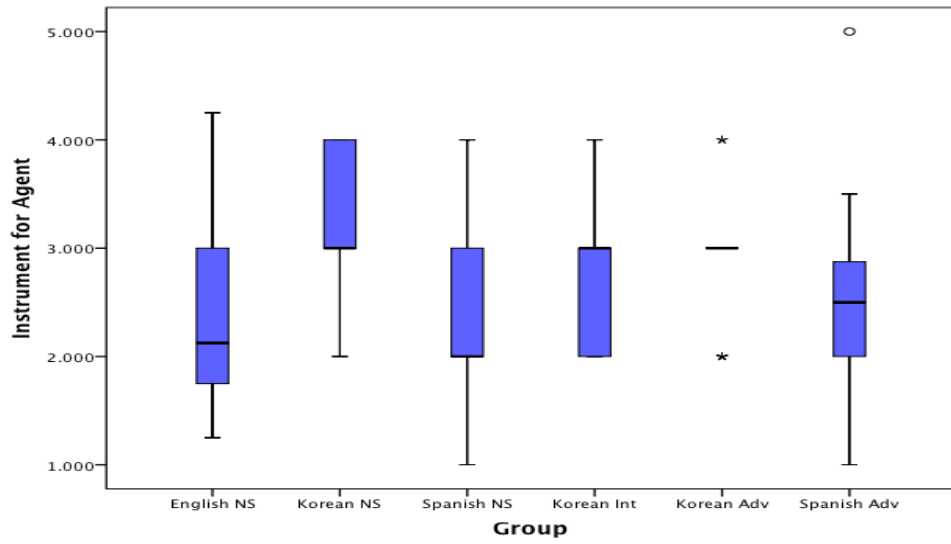
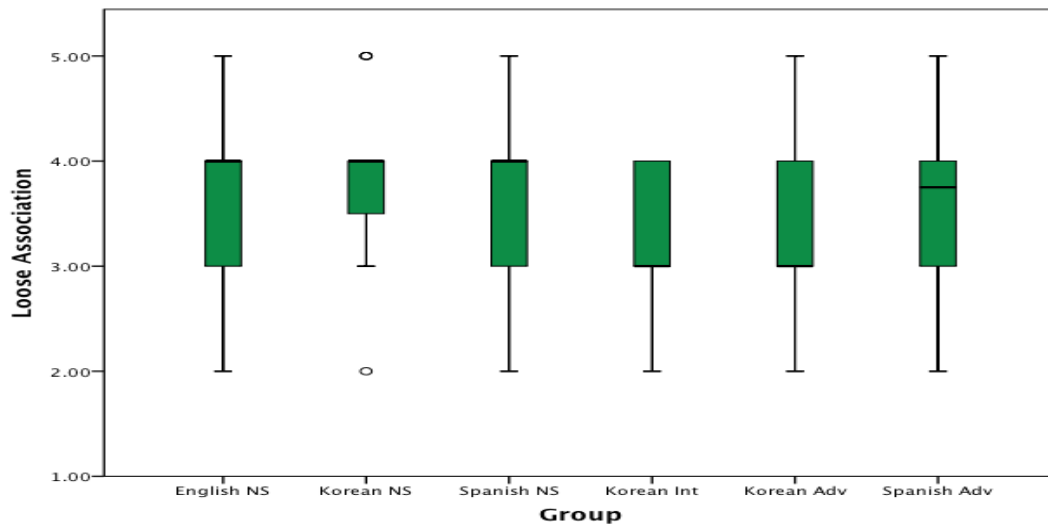Figure 3: Boxplots for the INSTRUMENT FOR AGENT metonymy in the Acceptability Judgment

Task

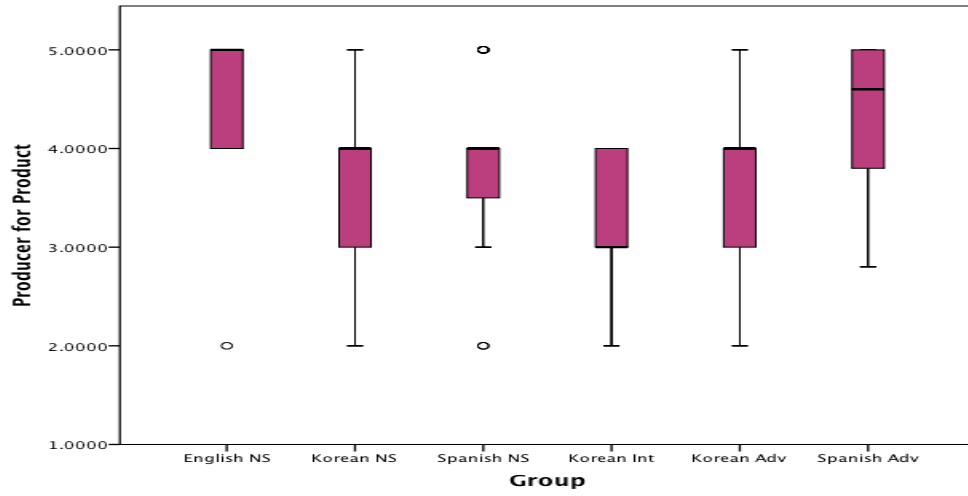Figure 4: Boxplots for the LOOSE ASSOCIATION metonymy in the Acceptability Judgment Task

Figure 5: Boxplots for the PRODUCT FOR PRODUCER metonymy in the Acceptability Judgment
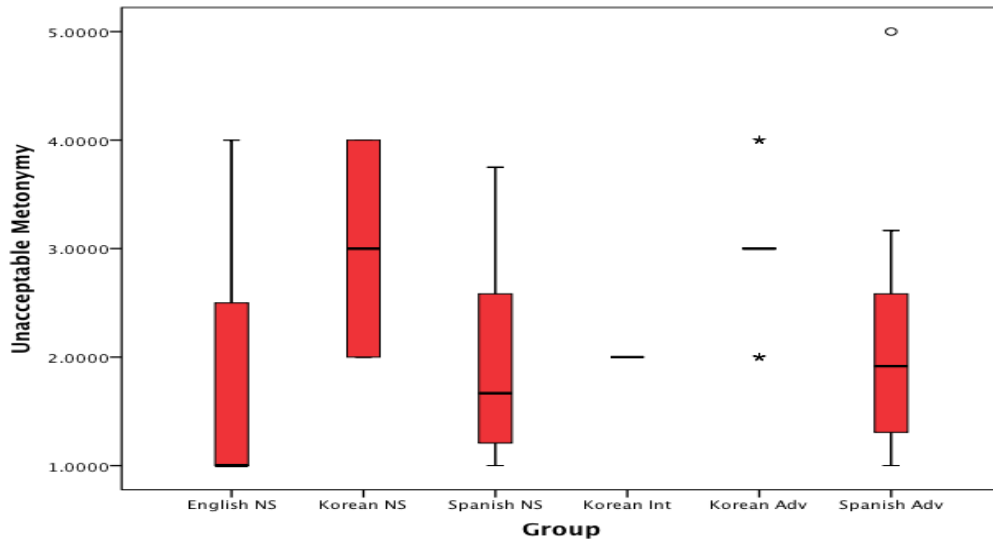
Task

Figure 6: Boxplots for the Unacceptable Metonymy condition in the Acceptability Judgment

Task

Endnotes:

[1] This experiment shows that even within the same metonymic pattern, PRODUCER FOR PRODUCT, particular instances can be novel, based on familiarity and frequency. Thus it seems that not only metonymic patterns but also particular instantiations of patterns can be located at different points on the conventionalization cline. We are grateful to an anonymous reviewer for making this point. We have not taken this fact into account in our research design, but it certainly deserves scrutiny with the competence of second language speakers, which we leave for further research.

[2] We thank two anonymous reviewers and the editor for raising this fair and important point.

[3] Another approach would have been to obtain naturalistic examples from Korean and Spanish in the patterns that we investigated. This would have worked well for the regular patterns but perhaps not so well for the unconventional patterns. We leave this approach for further research, and we thank two anonymous reviewers for suggesting it.

[4] The following is an example from a baseline sentence:
(i)   The painter is in a bad mood.
                    a. The worker who puts paint on walls is not happy today.
                    b. The brush that puts paint on walls is not happy today.
                    c. Both
                    d. Neither

[5] In addition, there could be a case of odd personification going on in this example, since the workbench is not clearly related to a person associated with it, but the bench itself is forced by the context to behave as a person.