The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Morphological learning in a novel language: A cross-language comparison

Viktória Havas[ab], Otto Waris[c], Lucía Vaquero[b], Antoni Rodríguez-Fornells[abd] & Matti Laine[c]
[a] Cognition and Brain Plasticity Group, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain
[b] Department of Basic Psychology, Campus Bellvitge, University of Barcelona, L'Hospitalet de Llobregat, Barcelona, Spain
[c] Department of Psychology and Logopedics, Abo Akademi University, Turku, Finland
[d] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
Accepted author version posted online: 14 Nov 2014.Published online: 09 Jan 2015.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2015
http://dx.doi.org/10.1080/17470218.2014.983531

Morphological learning in a novel language: A cross-language comparison

Viktória Havas[1,2], Otto Waris[3], Lucía Vaquero[2], Antoni Rodríguez-Fornells[1,2,4], and Matti Laine[3]

[1]Cognition and Brain Plasticity Group, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain
[2]Department of Basic Psychology, Campus Bellvitge, University of Barcelona, L'Hospitalet de Llobregat, Barcelona, Spain
[3]Department of Psychology and Logopedics, Abo Akademi University, Turku, Finland
[4]Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

Being able to extract and interpret the internal structure of complex word forms such as the English word *dance+r+s* is crucial for successful language learning. We examined whether the ability to extract morphological information during word learning is affected by the morphological features of one's native tongue. Spanish and Finnish adult participants performed a word–picture associative learning task in an artificial language where the target words included a suffix marking the gender of the corresponding animate object. The short exposure phase was followed by a word recognition task and a generalization task for the suffix. The participants' native tongues vary greatly in terms of morphological structure, leading to two opposing hypotheses. On the one hand, Spanish speakers may be more effective in identifying gender in a novel language because this feature is present in Spanish but not in Finnish. On the other hand, Finnish speakers may have an advantage as the abundance of bound morphemes in their language calls for continuous morphological decomposition. The results support the latter alternative, suggesting that lifelong experience on morphological decomposition provides an advantage in novel morphological learning.

Keywords: Morphology; Word learning; Second language acquisition; Cross-language differences.

One of the most important traits of human language is its unique ability to generate an infinite number of utterances from a limited set of particles. When we learn a language, may that be as infants acquiring our native tongue (L1), or as adults during second language (L2) acquisition, we need to learn the linguistic representations that carry semantic information (lexical units) and to discover the dependencies between these units (syntax). At the crossroads between lexicon and syntax lies morphology, the internal structure of words that a language learner has to master. Being able to decode bound morphemes is crucial as these elements can modulate the meaning of the word stem, the syntactic class of the word, and its thematic role in a sentence. Consider, for example, the English word form *build+ing+s* where the combination of these

Correspondence should be addressed to Matti Laine, Department of Psychology and Logopedics, Abo Akademi University, FI-20500 Turku, Finland. E-mail: matti.laine@abo.fi

M.L. was financially supported by the Academy of Finland [project number 260276]. This project has also been supported by the Catalan Government [grant number 2009SGR93]; and the Spanish Government (MICINN) [grant number PSI2011-29219] awarded to A.R.-F.

three morphemes defines the meaning of the word, the grammatical class of the word form, and the quantity. Despite the central role of morphological knowledge in language learning, experimental evidence on the acquisition of morphology is still rather limited (Endress & Hauser, 2011; Ferman & Karni, 2010; Ferman, Olshtain, Schechtman, & Karni, 2009; Merkx, Rastle, & Davis, 2011; Tamminen, Davis, Merkx, & Rastle, 2012).

In the present study, we examined the effects of the morphological structure of one's native language on the acquisition of a bound morpheme in a novel language. We compared two languages that differ considerably in terms of morphological structure: Spanish and Finnish. To study morphological learning in speakers of these two languages, we designed an associative word–picture learning task in an artificial language where the grammatical feature gender was embedded in the to-be-learned novel words as a bound morpheme. Spanish carries gender marking but has more limited bound morphology than Finnish, which, in turn, lacks gender but requires continuous morphological decomposition due to the abundance of bound morphemes.

Grammatical relations between words can be expressed in many ways. Some languages use fixed word order, others use particles (such as prepositions), and yet others use intonation (tonal languages, e.g., Chinese or Thai). A fourth mechanism used by a wide range of languages is an elaborate morphological system through which dependencies are reflected via bound morphemes added to the lexeme. Most languages use a combination of the different methods, which gives rise to a great diversity of morphological systems that are highly variable in their complexity (Comrie, 1989a; Pirkola, 2001; Whaley, 1997).

As regards the morphological processes of derivation and inflection, the former one carries robust semantic information, can modify word class, and have a restricted productivity (e.g., English *-tion*: satisfy–satisfaction). On the other hand, inflectional affixes chiefly provide grammatical information, they never alter word class, and their productivity is unrestricted (e.g., English *-s*: house–houses; Whaley, 1997). The focus of the present study is on the acquisition of a gender-marking system, which is mostly considered to belong to the realm of inflectional morphology, although it has certain derivational aspects as well, stemming from the origins of gender systems in Proto-Indo-European language (Luraghi, 2011). According to some linguists (e.g., Luraghi, 2011), in some modern Indo-European languages such as Romance languages, gender may still be in part derivational, although its derivational function is very limited.

Even though Spanish nouns have a rich derivational morphology, their inflectional structure is very limited. This is similar to nouns in most of Romance languages and results from the elimination of the declinational system of Latin (Comrie, 1989b). As described in Ambadiang (1999), there are two functions expressed in a Spanish noun: number (singular/plural), and gender (masculine/feminine). Most nouns have a fixed gender (either masculine or feminine), and the number is expressed by adding a suffix to the end of the word (a null morpheme for singular and an *-s* for plural). For example, the word "lámpara" (lamp) is a feminine noun singular. Combining it with the suffix *-s* yields the plural form "lámparas" (lamps), and thus these nouns have two possible forms. However, there are nouns that can have both masculine and feminine endings, generally depending on the natural gender of the agent the word denotes. Consider, for example, the word "oso" (bear). When talking about a male bear we use the word "oso" (word stem + male gender suffix), but if we talk about a female bear the word form is "osa" (word stem + female gender suffix). These nouns, therefore, have four possible forms: masculine-singular, masculine-plural, feminine-singular, and feminine-plural. The categories of number and gender are marked on determiners, demonstratives, pronouns, and adjectives, as well as nouns.

Gender represents a noun class system that is present in various forms in many languages (Corbett, 1991). Gender assignment in a language can be based, for example, on biological sex (male/female), human/nonhuman, or animate/inanimate classifications. Gender systems can be divided into semantic systems and formal systems. In the semantic systems, the meaning of a noun determines the gender category. In pronominal gender systems like English, for example, the pronoun *"he"* always refers to male humans or in some cases male domestic animals, and *"she"* always refers to female humans and female domestic animals. All other nouns fall into the neuter category and are referred to as *"it"*. In the formal systems, phonological or morphological properties of the noun divide them into different groups. For example, Italian nouns ending in *"-a"* are feminine (casa–house, stella–star), and those ending in "*-o*" are masculine (gelato–ice cream, soffitto–ceiling). Most languages with gender have partly semantic and partly formal systems. In Spanish, when the noun denotes a person or an animal, gender is marked according to the natural gender of the referent (e.g., profesor–male teacher, profesora–female teacher, león–male lion, leona–lioness). Nouns that represent inanimate objects are assigned a gender chiefly based on phonology. Words that end in *-a, -dad, -ción, -sión, -zón, -dez*, or *-iz*, for instance, are mostly feminine (e.g., manzana–apple, ciudad–city, razón–reason) with the exception of words ending in *-ma* that originate from Greek (e.g., fantasma–ghost, morfema–morpheme), which are always masculine. Nouns that end in *-o, -or,* or *-aje* are usually masculine (e.g., gato–male cat, colour–colour, viaje–journey).

Conversely, even though Finnish lacks the attribute gender, it has an extremely complex nominal morphology with 85 declinational classes, which yields over 2000 possible forms for each noun (Comrie, 1989b; Karlsson, 1983). Morphology of the Finnish noun (and adjective) provides for the expression of five functions, and therefore the noun has five slots in fixed sequence: (a) stem, which includes any embedded derivational suffix, (b) number (singular or plural), (c) case suffix, which expresses a wide variety of relations (e.g., temporal and spatial), (d) personal possessor suffixes, and (e) enclitic suffixes, which are used to express emphasis or subjective views and impressions (Comrie, 1989b).

To illustrate the difference in morphological structure between Spanish and Finnish, consider the following example:

1. a. ystäv-i-ltä-än-kö
   friend-P-DIR-POSS-INTER
   "from his/her friends, eh?"
   b. de su-s amig-o-s, eh?
   from his/her-P friend-MSC-P eh
   "from his/her friends, eh?"

This single example illustrates clearly how some of the information that is expressed via suffixes in Finnish is expressed with the use of prepositions and pronouns in Spanish.

Extracting word-internal morphological structure is analogous to another segmentation process during novel language learning that has been studied extensively—namely, word segmentation from a continuous speech stream. These studies have shown that both infants (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996) and adults (Saffran, Newport, & Aslin, 1996) can segment word candidates from continuous speech by using transitional probabilities of the syllable sequences. Moreover, participants are more prone to associate these "protowords" with external referents than other novel phoneme strings (e.g., Mirman, Magnuson, Estes, & Dixon, 2008). Following this, one would also assume that high-frequency phoneme/letter combinations embedded in otherwise highly variable novel words are more likely to be identified as potential bound morphemes in that language.

With regard to L1 learning, recent experimental evidence from infants suggests that bound morphemes begin to be extracted and recognized early on during development. Marquis and Shi (2012) found that 11-month-old preverbal infants learning French can identify a frequently occurring suffix in verbs and are able to learn and generalize even a completely novel suffix in an artificial language. In another study with infants learning English, Mintz (2013) showed that 15-month-olds decomposed novel words carrying the common suffix *-ing*. In summary, these studies indicate that infants start to develop mental representations for bound morphemes prior to using them in their own speech.

Concerning adults, Ferman et al. (2009) studied the acquisition of an artificial morphological rule superimposed on existing Hebrew verbs in a small group of young Hebrew-speaking adults. They employed a dependency rule where the novel suffix on a verb was governed by the animacy of the subject of a sentence. The training procedure spanned over three days, and the rule was never explicitly taught. The results indicated significant learning of the morphological rule as measured by both judgement and production tasks, with the participants being able to generalize the rule to new subject–verb pairs. A further study by the authors showed that young adults outperformed 8- and 12-year-old children on this training paradigm (Ferman & Karni, 2010). A language learning paradigm combining existing word stems with novel suffixes was employed by Merkx et al. (2011). In their study, English-speaking participants were exposed to novel bound morphemes attached to familiar word stems (e.g., leap+esh) and presented either with or without a semantic definition (e.g., "The cost of having a stuntman leap out of a building"). Their results indicated that irrespective of the provision of the definition, the novel suffix was successfully extracted by the participants. At the same time, the semantic condition increased false recognition of new items carrying the trained suffixes, suggesting that suffixes presented with semantic information were in the process of becoming lexicalized. Tamminen et al. (2012) successfully used the same paradigm to provide evidence for the generalization of the newly learned suffixes. Finally, using a simple artificial grammar learning paradigm where novel stems were presented auditorily with an affix-like element /Zal/, Endress and Hauser (2011) examined the effects of type and token frequency on subsequent recognition and generalization performance. Their results indicated that high token frequencies are important for the memorization of exceptions (e.g., items with string-initial /Zal/ when most items in a set carry it at the end) while sufficient type frequency is needed to yield generalization of the rule to totally new strings.

The experimental studies briefly reviewed above suggest that adults are effective in extracting and generalizing morphological information from novel polymorphemic words without explicit instruction. However, the previous learning paradigms have either addressed more general forms of learning via syllable frequency and position effects (Endress & Hauser, 2011) or combined existing word stems with novel suffixes (Ferman & Karni, 2010; Ferman et al., 2009; Merkx et al., 2011) that provide strong cues for initial identification of the new morphemic units. In our study, we introduced an artificial language learning paradigm that is closer to L2 learning by exposing the learner to totally new, semantically meaningful word stems and suffixes. Our paradigm was a simple word–picture learning task where the target words carried a single grammatical feature—namely, semantic gender signalled by a bound morpheme. We chose semantic gender as it can be readily signalled in the pictured referents. Moreover, gender provides an interesting contrast between the two languages we studied as it is present in Spanish but not in Finnish.

One aspect of L2 learning that has been highlighted in earlier studies is the importance of the structural relationships between L1 and L2 for the acquisition of L2. With regard to gender, Sabourin, Stowe, and de Haan (2006) showed that L2 acquisition of grammatical gender is affected by the degree of similarity of the gender classification systems between L1 and L2. Their participants were Dutch L2 learners whose mother tongue was German (a language very similar in its gender system to Dutch), a Romance language (French, Italian, or Spanish that have a semantic and phonological gender system but with a very different gender agreement pattern from Dutch), or English (a language that only has a pronominal gender system). When testing for acquisition of gender agreement in Dutch, significant group differences emerged with the German group performing best, the Romance group being above chance, and the English group remaining at chance level.

For the present study, we created an artificial language with an overt marking for semantic gender. This system bears similarity to the gender classification system of Spanish and Catalan, the native languages of our Spanish participants. In contrast, Finnish, a non-Indo-European agglutinative language, lacks gender marking. However, it has a very rich inflectional system that yields over 2000 possible forms for each noun and over 10,000 forms for each verb (Karlsson, 1983). The huge number of inflected word forms encountered by speakers of Finnish calls for frequent online morphological decomposition, as has been repeatedly found in experimental studies on word recognition in Finnish (e.g., Hyönä, Laine, & Niemi, 1995; Laine, Vainio, & Hyönä, 1999; Soveri, Lehtonen, & Laine, 2007).

The L1 differences outlined above lead to two opposing hypotheses. Given the results of Sabourin et al. (2006), one could hypothesize that the Spanish participants will be better at acquiring the gender rule of our artificial language, as it is similar to the gender structure in their native languages. On the other hand, while the Finnish learners lack gender in their L1, they may be more prone to morphologically decompose novel items that may provide a general advantage in morphological learning in a new language. Finally, at least under certain circumstances the existence of a particular gender structure in one language might interfere with the acquisition of another gender system. It has been observed that gender mistakes are quite common amongst second-language learners when acquiring words that do not match in gender across languages (e.g., Spanish and German; see, for example, Rodríguez-Fornells, de Diego Balaguer, & Münte, 2006). In the first experiment, we compare a Spanish- and a Finnish-speaking group to evaluate our morphological learning paradigm and the influence of L1 on second language learning. This is followed by a second experiment that attempts to replicate the first one for the Spanish-speaking participants and evaluates a possible phonological interference from Spanish L1 on the acquisition of the gender system of our artificial language.

EXPERIMENT 1

Method

Participants

To assess the influence of the first language on morphological learning, two groups of healthy participants were tested: (a) a group of 31 Finnish-speaking university students (22 women, $M_{age} = 23.91$ years, $SD_{age} = 4.05$ years), and (b) a group of 59 Spanish-speaking university students (39 women, $M_{age} = 21.10$ years, $SD_{age} = 4.17$ years). None of the participants had neurological problems or diagnosed learning disabilities such as dyslexia. They all had normal or corrected-to-normal vision. No group differences were found for gender distribution, $\chi^2(1) = 0.11$, $p = .26$, or on the scaled scores of the Similarities subscale of the Wechsler Adult Intelligence Scales–III (WAIS–III; Wechsler, 1997), which was used as a measure of verbal intelligence, $t(89) = -0.561$, $p = .576$. However, we did find differences concerning foreign language acquisition. All participants except one had previous experience with one or more foreign languages. Finnish participants typically had learned Swedish and English in addition to other languages; Spanish participants usually had learned English and a Romance language other than their mother tongue and, in some cases, also other languages like German. We compared the number of languages that participants spoke; Finnish participants had more language learning experience, $t(89) = -5.94$, $p < .0001$ (Finnish: $M = 4.53$, $SE = 0.18$, $SD = 0.98$; Spanish: $M = 3.41$, $SE = 0.10$, $SD = 0.79$). To address this issue, we included the number of spoken languages as a covariate in our analyses.

Stimuli

The to-be-learned training stimuli consisted of visually presented word–picture pairs (WPPs). The cartoon-like black-and-white pictures depicted both living and nonliving objects (see Figure 1). Part of the pictures (targets) depicted animals that had stereotypical Western male or female clothing/appearance while others were neutral in terms of gender (fillers). The picture

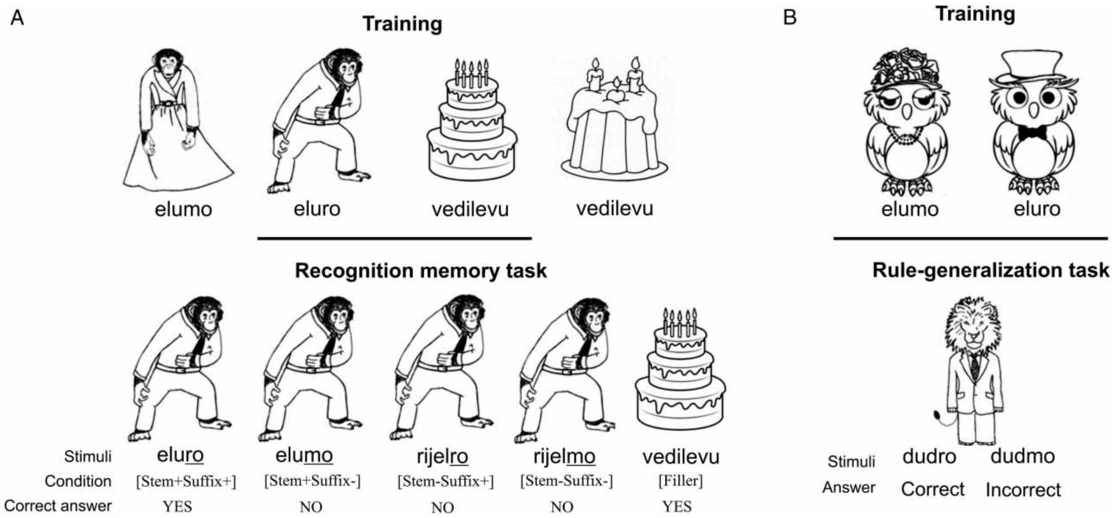

Figure 1. *A. Example for the experimental conditions of the recognition memory task. B. Examples of the stimuli used in the rule generation task.*

names were pronounceable novel strings varying in length from 5 to 8 letters. In addition, the target names carried a gender marking in the form of a suffix. Two suffix pairs (ro-so, ga-sa in the case of the Finnish participants and mo-ro, za-ga for the Spanish participants) were employed, with both the pairs and their gender assignment counterbalanced across participants. For each object, two different pictures appeared in the training set. For the targets, these were the male and female variants of the same animal, and for the fillers visually somewhat different renditions of the same object (see Figure 1).

Altogether 64 WPPs were created as training material. They consisted of 32 gender-marked WPPs (16 different animal species) and 32 fillers (16 pairs with two images per referent). To introduce variability to the novel strings, the word stems were constructed so that four stems had three letters (CVV, CVC, VVC, & VCV; C = consonant, V = vowel), four stems had four letters (CVCV, VCVC, VCCV, & VCCV), four stems had five letters (CVCVC, CVCVC, CVCCV, & CVCCV), and four stems had six letters (CVCVCV, VCVCVV, VCVCVC, & CVCCVC). The novel words presented to the Spanish versus Finnish participants differed slightly so that they were phonotactically legal and consisted of syllables that existed in the participants' mother tongue (e.g., "zavepro" in the Spanish vs. "vavepro" in the Finnish version).

Experimental procedure

The experiment was completed during a single session that lasted for about two hours. The session included three tasks and a short structured interview in addition to the WAIS–III Similarities test and a questionnaire.

Training. During the training phase, the participants were simply instructed to memorize as many of the 64 WPPs as possible. The WPPs were presented on a computer screen against a grey background with the words written in black. Each WPP was shown for 3500 ms with a 500-ms blank interval between the WPPs. The complete list of the 64 WPPs was presented eight times, with a brief pause after every 16 WPPs. The presentation order of the WPPs was randomized for each presentation round, separately for each participant. The training task took about 25 minutes.

To eliminate the possibility that any of the WPPs would be retrieved from short-term memory, both groups performed a short cognitive task prior to being tested for their learning performance. The Finnish participants performed an auditory nonword span task where they listened to series of artificial words (different from the trained words) and were asked to memorize them. After hearing each series of words, they wrote the words on an answer sheet. The test started with a three-word series, and the number of words increased one by one for each series. The Spanish participants performed a segmentation task; they saw several Spanish words on the screen presented as a letter string without space between them. They were asked to count the number of words presented on the screen at once and type in that number. Both tasks took between 5 and 10 min.

Recognition memory task. In this task, the participants were presented with WPPs, and they were to press the left mouse button if the word and the picture matched and the right mouse button if the word and the picture did not match. In a factorial design, we devised stimuli where (a) both the stem and the suffix matched with the picture, (b) only the stem or (c) only the suffix matched, or (d) neither morpheme matched. In addition, we included filler items in the stimuli. Examples for the different conditions can be seen in Figure 1A. The purpose of the setup was twofold. First, the participants' ability to separate fully correct WPPs versus fully incorrect WPPs gave a measure of overall word learning ability that did not hinge upon morphological learning. Second, the participants' performance on the different types of incorrect WPPs was expected to show whether morphological learning had taken place. Previous evidence from a similar word–picture matching task with familiar word forms indicates that the word stem is the primary element in the meaning analysis of a suffixed word (Laine, 1999). Accordingly, of particular interest was the participants' performance on items where the word stem matched the picture but the gender suffix did not. A participant who had learned the word but failed to acquire the meaning of the gender suffix would rely on the primary element (i.e., the stem) and thus be prone to incorrectly reply "yes" to such an item. In contrast, a participant who had even learned the meaning of the gender suffix should be able to reject such an item, albeit with a longer decision latency as the saliency of the matching stem makes the decision more difficult (Laine, 1999).

The Spanish version of the task comprised 48 trials divided into the five trial types: gender-marked pictures with correct stem and correct ending (stem+suffix+; $n = 8$), gender-marked pictures with correct stem and incorrect ending (stem+suffix–; $n = 8$), gender-marked pictures with incorrect stem and correct ending (stem–suffix+; $n = 8$), gender-marked pictures with incorrect stem and incorrect ending (stem–suffix–; $n = 8$), filler pictures with correct names ($n = 16$; see Figure 2). This gave a total of 24 "yes" (left mouse button) responses and 24 "no" (right mouse button) responses. The Finnish version was slightly different, with additional 8 incorrect and 8 correct fillers, yielding a total of 64 trials (32 "yes", and 32 "no" answers). The number and types of the target items were identical to those in the Spanish version. Within both test versions, the average bigram frequencies between the targets and the fillers versus the foils were comparable [Spanish version: $F(1, 141) = 0.511, p = .601$; Finnish version: $F(1, 141) = 0.112, p = .895$]. The incorrect gender marking always denoted the opposite sex. This way all pictures, stems, and suffixes presented were part of the training stimuli. All stimulus groups were also counterbalanced regarding the specific WPPs that were correct/incorrect in the task.

Each WPP was presented on the computer screen until the participant pressed a mouse button or until 3500 ms had passed. A trial was classified as an omission if the participant did not respond within the time limit. All participants were made aware of this time limitation before beginning the task. A 500-ms blank interval appeared after each WPP.

Rule generalization task. The present task evaluated the participants' ability to generalize the novel gender-marking system to new stems. The participants were presented with completely new pictures and letter strings so that each picture was

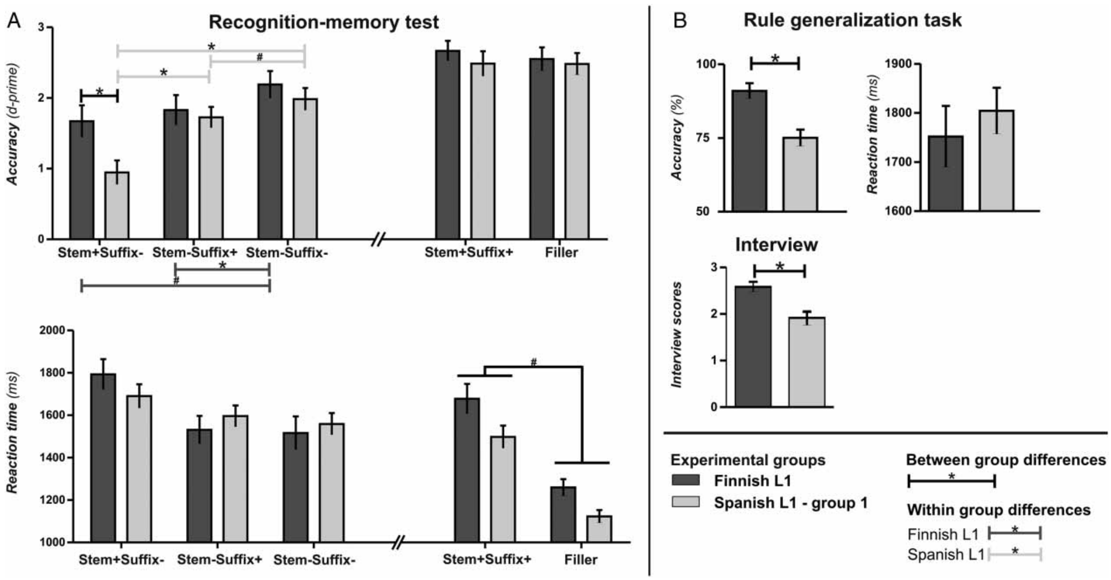

Figure 2. *A. Average* d′ *values and reaction time (RT) of the conditions of the word recognition memory task for the Spanish and Finnish groups. L1 = native language. Statistically significant pairwise differences between conditions in the Finnish-speaking group are marked in dark grey, between conditions in the Spanish group in light grey, and between groups in black. B. Accuracy and RT of the rule-generalization tasks and scores of the interview are depicted.* *$p < .01$. #$p < .08$. *Error bars represent the standard error of the mean.*

coupled with two strings that had the same (previously unseen) stem but a different gender-marking suffix that was familiar from the training stimuli (see Figure 1). These novel items consisted of 32 gender-marked pictures depicting 16 animal species. By pressing the corresponding button, the participant was to choose which of the two words (left or right) better matched the picture. Participants had 3500 ms to respond, and a 500-ms blank interval separated the trials. The stems for the new items were constructed in the same way as the stems used in the learning task and the word-picture matching task (see above).

Interview. After the completion of the rule generalization task, the participants were asked a few short questions concerning their explicit knowledge of the gender-marking system embedded in the WPPs. The participants scored from −1 to 3 depending on their explicit awareness of the gender-marking feature. To get the highest score of 3, the participant had to spontaneously report the gender-marking feature on the general question "*How would you describe this language*?" If the participant did not report the feature, the second question "*Have you noticed any regularities in this new language?*" was asked, and the participant was awarded 2 points if the gender-marking system was described at this point. If the participant still did not report the morphological feature, the third question "*Were certain words, or parts of words, more common than others*?" was asked. At this point, the participant was awarded 1 point for reporting the gender-marking system. If the participant did not report the specific word endings after the three questions, a sheet with six consonant–vowel pairs (Finnish: SO, GA, SA, PE, RO, TI; Spanish: MO, GA, ZA, PE, RO, TI) was presented, and the participant was asked to point out which of the consonant–vowel pair/pairs they thought were the most common endings in the training material. For each correct response, 0.25 points were awarded, while 0.25 points were subtracted for every incorrect response. Thus the total score ranged between −1 and 3 points.

Background test. WAIS–III Similarities. The Similarities subscale of the WAIS–III intelligence test was administered to obtain an estimate of the verbal IQ of participants. As described above, no significant difference was found between the groups on this measure.

Background questionnaire. A short questionnaire was administered to both groups regarding education, occupation, known and studied languages, vision, hearing, past or present problems with reading and/or writing, possible neurological and psychiatric diagnoses, present medications, and/or other drugs.

Results

Learning outcomes in the two groups

We sought for evidence for group differences in learning by performing mixed-model analyses of covariance (ANCOVAs)—with task condition and L1 as factors and foreign language experience (FLE) as covariate—on target detection as expressed by d' values and on reaction times (RTs) of the recognition memory test, on the accuracy rates and reaction times in the rule generalization test, and on the interview scores. Variance inhomogeneities were evaluated, and corrected values are reported when necessary. The RT analyses included only correct trials. Furthermore, RTs that were below 500 ms or two standard deviations above the individual mean RT were also excluded.

Recognition memory task. At first, we conducted an analysis of d' values to assess whole word learning in the two groups. To that end, we contrasted the fillers and the stem+suffix+ condition separately against the stem–suffix– condition (the accuracy values in percentages of all the conditions can be found in Table 1), as these are the conditions where participants could respond correctly even in the absence of any morphological knowledge of the trained words. Mixed-model ANCOVA showed no significant main effects or interaction [condition: $F(1, 87) = 0.029$, $p = .865$; L1: $F(1, 87) = 0.059$, $p = .808$; FLE: $F(1, 87) = 0.144$, $p = .706$; Condition × L1: $F(1, 87) = 0.224$, $p = .637$; Condition × FLE: $F(1, 87) = 0.129$, $p = .720$]. The d' values are depicted in Figure 2A.

To evaluate morphological learning in the two groups, mixed-model ANCOVA with the within-group factor condition (3 levels: stem+suffix–, stem–suffix+, stem–suffix–) and the between-group factor L1 (2 levels: Finnish, Spanish) was calculated for both d' values and RTs. We calculated d' values for the three untrained conditions using accuracy scores of the stem+suffix+ condition as hits, and the incorrect "yes" responses in the three conditions as false alarms. We found a main effect of L1, reflecting the higher overall accuracy rate of the Finnish group, $F(1, 87) = 4.939$, $p = .029$, and no main effect of condition, $F(2, 174) = 0.386$, $p = .680$, or FLE, $F(1, 87) = 0.036$, $p = .851$. Moreover, the Condition × L1 interaction was significant, $F(2, 174) = 5.156$, $p = .007$, stemming from the fact that Finnish-speakers were better than Spanish-speakers at the stem+suffix– condition, $p = .001$, whereas no group differences on the other two conditions were found on post hoc analyses, $p_{\text{stem}-\text{suffix}+} = .434$, $p_{\text{stem}-\text{suffix}-} = .178$. Additionally, we found a significant difference between the stem+suffix– and stem–suffix+ conditions ($p < .0001$) in the Spanish but not in the Finnish group. The d' values and significant differences are depicted in Figure 2A. The Condition × FLE interaction term remained nonsignificant, $F(2, 174) = 0.524$, $p = .368$.

A mixed-model ANCOVA on the RTs did not show any significant effect for condition [$F(2, 152) = 0.127$, $p = .880$; L1: $F(1, 76) = 0.943$, $p = .335$] or for the Condition × L1, $F(2, 152) = 0.722$, $p = .487$, and Condition × FLE interactions, $F(2, 152) = 1.913$, $p = .151$. The main effect of FLE did not quite reach statistical significance, $F(1, 76) = 3.255$, $p = .075$.[1]

Rule generalization task. One-sample t-tests showed that both groups evidenced morphological learning by performing above chance level in this

[1]Note that by-item analyses were not performed as the stimulus sets were not identical for the two groups.

Table 1. *Mean accuracy rates, the standard error of the mean, and the standard deviation of the five experimental conditions of the recognition-memory test in Experiments 1 and 2*

Group	n	*Stem+suffix+*			*Stem+suffix−*			*Stem−suffix+*			*Stem−suffix−*			*Filler*		
		M	SEM	SD	M	SEM	SD	M	SEM	SD	M	SEM	SD	M	SEM	SD
Finnish group	31	.86	.02	.09	.73	.05	.27	.78	.04	.25	.90	.02	.12	.84	.02	.09
Spanish group 1	59	.84	.03	.20	.46	.04	.34	.73	.03	.24	.81	.03	.23	.87	.02	.17
Spanish group 2	38	.80	.04	.24	.56	.06	.35	.71	.04	.27	.74	.05	.30	.79	.04	.22

Note: Experiment 1: Finnish group, Spanish group 1; Experiment 2: Spanish group 2. Accuracy rates in percentages.

task [Finnish: $t(31) = 15.30$, $p < .0001$; Spanish: $t(58) = 8.65$, $p < .0001$]. As Figure 2B shows, the Finnish participants had a higher average rate of accuracy than the Spanish participants. Univariate ANCOVA showed that this group difference was statistically reliable, $F(1, 88) = 12.40$, $p = .001$. We found no effect of FLE, $F(1, 88) = 0.75$, $p = .389$. There was a marginally significant difference in response latencies between the two groups, $F(1, 88) = 3.20$, $p = .077$, with the Finnish participants being somewhat faster than the Spanish participants. We also found a positive effect of FLE on response latency, $F(1, 88) = 0.267$, $p = .025$.

Interview. Univariate ANCOVA showed that the Finnish group reached higher scores on the interview, $F(1, 88) = 6.878$, $p = .01$ (see Figure 3B), indicating that they were more aware of the morphological rule than the Spanish participants. We found no effect of FLE, $F(1, 88) = 0.27$, $p = .607$. To assess the relationship between performance on the recognition task and the interview, we calculated the correlations between the d' rates on the recognition memory task and the interview scores. For the stem+suffix− condition, this correlation was found to be statistically significant, $r(88) = .425$, $p < .0001$, indicating that higher accuracy on this condition was associated with better subjective awareness of the morphological rule. The correlations for the stem−suffix+ condition, $r(88) = -.020$, $p = .851$, and the stem−suffix− condition, $r(88) = .155$, $p = .146$, were nonsignificant. We also found a positive correlation between the interview scores and the accuracy rates on the rule-generalization test, $r(88) = .676$, $p < .0001$.

Foreign language experience

We evaluated the participants' previous FLE and related it to their performance on the experimental tasks. When we pooled the data of the two groups together, a statistically significant correlation emerged between FLE and interview score [interview score: $r(89) = .226$, $p = .031$; stem+suffix− d': $r(89) = .102$, $p = .341$; rule-generalization accuracy: $r(89) = .118$, $p = .264$]. However, when

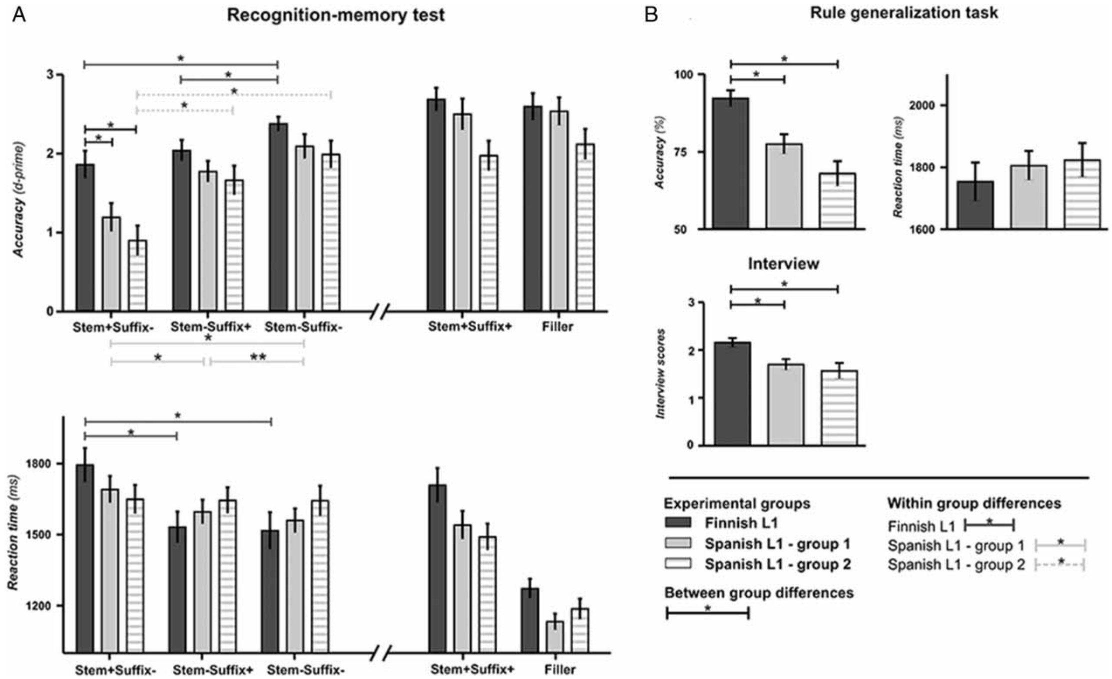

Figure 3. *A. Average* d′ *and reaction time (RT) values of the conditions of the recognition memory task for the groups of Experiments 1 and 2. L1 = native language. B. Accuracy and RT of the rule-generalization tasks and scores of the interview for all three groups of Experiments 1 and 2 are depicted.* *p *< .01. Error bars represent the standard error of the mean.*

analysing the two groups separately, no significant associations between FLE and the results on the experimental tasks were found: Spanish group [interview score, $r(57) = .067$, $p = .616$; stem+suffix– d', $r(57) = -.170$, $p = .199$; rule-generalization accuracy, $r(57) = -.031$, $p = .816$]; Finnish group [interview score, $r(30) = .036$, $p = .844$; stem+suffix– d': $r(30) = .08$, $p = .668$; rule-generalization accuracy: $r(30) = -.239$, $p = .188$]. Importantly, statistically significant intercorrelations between the results of the experimental tasks and conditions remained even when the two groups were analysed separately. Moreover, we found no statistically significant effects of FLE in the ANCOVAs performed on the recognition-memory task results. The only statistically significant effect of FLE surfaced up in the response latencies of the rule-generalization task. These results suggest that the correlations in the pooled analysis were driven by the group differences in morphological learning rather than by FLE as such.

Discussion

We compared two groups of adult participants with different native tongues on a new morphological learning task. The aim was to test which of two potential factors, L1–L2 compatibility of the to-be-learned grammatical feature or the overall experience on real-life morphological decomposition, is more effective in facilitating the acquisition of a grammatical feature in an artificial language.

First, the results showed that on the recognition memory task, the Spanish and Finnish groups evidenced similar levels of overall word acquisition. This helps to rule out general word learning ability as a source of differential morphological learning between the two groups. Secondly, the groups differed significantly on all three measures of morphological learning that we employed. The Finnish participants showed more sensitivity to morphological structure in the word recognition task, as they rejected more accurately the particularly demanding word–picture pairs that carried the correct stem but the wrong suffix. To be able to do this, one must have acquired the meaning of the gender suffix. Moreover, the Finnish participants had a higher accuracy rate on the gender rule generalization task and showed more awareness of the embedded gender rule in the posttest interview. In other words, the participants with more experience in morphological decomposition had an advantage in acquiring the covert gender rule even though their native language lacked this grammatical feature. Therefore, our results suggest that for unsupervised learning of a novel grammatical feature, familiarity with a similar rule system in one's L1 is less influential than lifelong extensive practice with morphological decomposition required by L1. This was the case even though our Spanish participants had in fact a richer L1 experience in the early childhood by being Spanish–Catalan bilinguals; Finnish participants had mastered Finnish till they started formal education in primary school. Previous studies comparing monolinguals and bilinguals on word learning have shown an advantage for bilinguals especially in the case of concrete referents (Kaushanskaya & Marian, 2009; Kaushanskaya & Rechtzigel, 2012). However, our study did not only investigate word-referent mapping but morpheme segmentation within a word, and both of the languages of our bilingual participants (Spanish, Catalan) are morphologically more limited than the native tongue of our monolingual participants (Finnish).

With regard to the effects of L1 on grammatical gender processing in L2, Sabourin et al. (2006) studied this issue with native speakers of English, Romance languages, and German, who were learning Dutch. They tested gender agreement in noun–relative pronoun sequences embedded within sentences through a grammaticality judgement task. The German group outperformed both the Romance and the English groups; the Romance group performed above chance level; however, the English group did not. Based on these results the authors argued that L2 acquisition of grammatical gender is affected by the morphological similarity of gender marking in the L1 and L2. Our results seem to contradict these findings. However, there can be several reasons for this discrepancy. First, the overlap in gender assignment between German and Dutch provided an advantage for the German L1 speakers. Second, the most clear-cut group differences were observed in a different gender feature from that in our study—namely, gender agreement with a noun and a relative pronoun. Third, in terms of overall morphological richness, especially English but also Romance languages are more limited than German. All these three factors could have accounted for the advantage of the German participants in their L2 acquisition of Dutch and its gender system. Finally, our gender assignment system could in principle have created some interference for our Spanish participants as not all the animal names depicted by our stimuli follow a similar morphological gender pattern to that of Spanish. In our artificial language all the animal names that carry gender marking consist of a stem and a gender-marking suffix; nevertheless, in Spanish some of the animal names are regular (as in the experimental language), for example: "elefant-e" (male elephant)–"elefant-a (female elephant)" or "gat-o" (male cat)–"gat-a" (female cat), but others follow a different pattern: One of the animals has a different morpheme added for female gender marking: "gall-o" (rooster)–"gall-ina" (hen), and in 10 of the animal species that we use, the corresponding Spanish word is either masculine or feminine (e.g., "chimpance" MASCULINE–chimpanzee, "hipopótamo" MASCULINE–hippopotamus, "tortuga" FEMININE–turtle).

There might also be a phonological interference from Spanish, as in Spanish language the word final -o usually marks masculine, and the word final -a mostly marks feminine gender. In our experiment, half of the participants had the suffixes -mo/-ro, and half of the participants had -za/-ga. Even though it is the consonant of the last syllable that carries the gender information, the word-final vowel could have made it more difficult for the Spanish participants to identify a noun ending in -a as masculine and a noun ending in -o as feminine. To investigate whether such interference exists, we ran a second experiment with slightly modified stimuli and a new group of Spanish participants.

EXPERIMENT 2

Here we tested a group of Spanish speakers with the same paradigm as that in Experiment 1, this time with modified stimuli to eliminate a possible gender-related phonological interference effect between Spanish and the artificial language used in the first experiment.

Method

Participants

A total of 38 (20 women) healthy university students from Barcelona between the ages of 18 and 49 years ($M = 24.43$, $SD = 5.787$) were recruited for the experiment. Compared to the Spanish participants in Experiment 1, there was no difference on the scaled scores of the Similarities subscale of the WAIS–III that was used as a measure of verbal intelligence; $F(2, 126) = 0.708$, $p = .49$.

Stimuli and procedure

We used the same experimental procedure and stimuli as those with the previous Spanish group, with one important difference: We substituted the suffixes -mo/-ro and –za/-ga with -mi/-ri and -zu/-gu. As in Spanish most of the nouns ending in -o are masculine, and those ending in -a are feminine, using the word-final -o and -a could have interfered with the Spanish participants' morphological learning process in Experiment 1. This potential interference should be effectively eliminated with the word-final vowels -i or -u, which are rarely used in that position in Spanish nouns.

Results

Word learning

We conducted a mixed-model ANCOVA with the factors: group (L1 Finnish, L1 Spanish 1, L1 Spanish 2) and condition (stem+suffix+, filler) on the d' values, with number of foreign languages studied (NoL) as a covariate. We used the stem–suffix– condition for false-alarm rates of the recognition-memory task. We found no within- or between-group differences [condition, $F(1, 124) = 0.101$, $p = .751$; group, $F(2, 124) = 2.286$, $p = .106$; Group × Condition, $F(2, 124) = 0.877$, $p = .418$]. We found no effect of NoL, $F(1, 124) = 0.652$, $p = .421$, or NoL × Condition interaction, $F(1, 124) = 0.108$, $p = .743$. The average d' values are presented in Figure 3A.

Recognition-memory test

The mixed-model ANCOVA on the d' values of the stem+suffix–, stem–suffix+, and stem–suffix– conditions showed no main effect of condition, $F(2, 248) = 0.954$, $p = .366$, a marginal effect of group, $F(2, 124) = 2.493$, $p = .087$, and a statistically significant interaction between the two factors, $F(4, 248) = 3.398$, $p = .019$. We found no effect of NoL, $F(1, 124) = 0.140$, $p = .709$, or NoL × Condition interaction, $F(2, 248) = 0.684$, $p = .467$.

Critically, the post hoc pairwise analysis showed no difference between the two Spanish groups in the stem+suffix– condition ($p > .99$); however, it did show a difference between the Finnish and Spanish 1 ($p = .006$) and the Finnish and Spanish 2 ($p = .003$) groups in the same condition. There was no difference between the groups in the stem–suffix+ or the stem–suffix– conditions ($p > .6$ in all comparisons; see the d' values in Figure 3A).

The same analyses on the reaction time results showed no main effect of condition, $F(2, 248) = 0.147$, $p = .842$, or group, $F(2, 124) = 0.298$, $p = .743$, but there was a statistically significant interaction between the two factors, $F(4, 248)$, $p = .05$. There was no effect of NoL, $F(1, 124) = 1.385$, $p = .242$, or a significant NoL × Condition interaction, $F(4, 248) = 1.366$, $p = .257$. As Figure 3A shows, the interaction stems from the following pattern: Finnish participants were significantly slower in the stem+suffix– condition than in both the stem–suffix+ ($p = .006$) and the stem–suffix– ($p = .002$) conditions; on the other hand, in the Spanish 1 group the stem+suffix– condition was significantly slower only when compared to the stem–suffix– condition ($p = .01$), while the comparisons did not reach significance ($p > .09$). Finally, the Spanish 2 group did not show any significant RT differences between these conditions ($p > .9$).

Rule generalization task

A one-way analysis of variance (ANOVA) on accuracy scores found a significant main effect of group, $F(2, 126) = 11.624$, $p < .0001$. Figure 3B shows that the Finnish participants had a higher accuracy rate than the other two groups ($p < .002$), and there was no significant difference between the two Spanish groups ($p = .210$). We found no group difference in the reaction times of this task, $F(2, 126) = 0.372$, $p = .690$.

Interview

The interview scores (see Figure 3B) revealed a similar picture to that for the accuracy rates for the rule generalization task. The one-way ANOVA showed a significant group effect, $F(2, 126) = 8.330$, $p < .0001$, with the Finnish participants having higher scores than the Spanish groups ($p < .01$). There was no significant difference between the two Spanish groups ($p = .695$).

Discussion

The second experiment tested whether a possible phonological interference effect of L1 (overlap between our artificial suffixes and Spanish gender-related word-final vowels) could have explained the Spanish participants' lower rates of morphological learning in the first experiment. Despite the elimination of this possible source of interference, we effectively replicated the results of the first experiment of the Spanish group in the recognition-memory test, the rule generalization task, and the interview. The two Spanish groups showed no difference on their accuracy rates in the critical stem+suffix– condition, while both groups performed worse than the Finnish group.

GENERAL DISCUSSION

In two experiments, we probed the initial stages of morphological learning in a second language. The main aim was to compare learning outcomes in speakers of structurally different languages, Finnish and Spanish. This comparison gave us the opportunity to test L1 influence on morphological learning under experimental conditions. The task we devised for this purpose was an associative word–picture learning task where we embedded a bound morpheme (suffix) in the to-be-learned words. Morphological learning was implicit as no explicit information was provided for this embedded feature.

The results of the two experiments can be summarized as follows: Experiment 1 showed that even though both Finnish and Spanish participants were equally successful in learning the novel words of the artificial language, the Finnish group was significantly better at learning the hidden gender rule and at generalizing it to new words and concepts. To control for possible phonological interference from Spanish gender marking, we tested a second group of Spanish participants. This second Spanish group showed a similar learning pattern to that of the first one. Their overall word learning was successful but they struggled learning the morphological rule even more than the first Spanish group.

While the present study was not designed to test general models of morphological learning, the two general views should be mentioned here (for a review, see Merkx et al., 2011). Form-based theories emphasize the sequential probabilities of letter combinations as a cue to detect morphemic boundaries in words, or the importance of the frequent appearance and combinatorial nature of affix-marking letter strings for their identification (see also Endress & Hauser, 2011). Semantically based theories, on the other hand, claim that form–meaning relationships are important in facilitating lower level orthographic learning of affixes. In fact, Merkx et al. (2011) found that semantic information of the novel suffix was necessary for lexicalization to take place. The learning paradigm we present here combines a regular gender-marking system with an unambiguous semantic content (male/female). It thus utilizes both form and meaning to provide a word learning context reminiscent of L2 acquisition. Further studies are needed to explore to what extent the learning advantage of the Finnish learners was related to a more effective initial extraction of the potential suffix at the form level, a better utilization of the form–meaning relationships, or both.

Some limitations of the present study are worth noting here. Our participant groups were comparable on their general background (young university students), sex distribution, performance on a verbal intelligence measure, and overall (nonmorphological) word learning ability in the particular task we used. However, as our study represents a natural groups design, we cannot rule out the possibility that the Spanish and Finnish groups differed in factors other than the nature of their mother tongue that contributed to the observed differences on morphological learning. Thus, the present results should be confirmed with new experiments using other language pairs. Another limitation deals with our measurement of awareness of the gender rule through the structured interview. As the interview was always the last task, it may have been confounded by the recognition and the generalization tasks. In other words, those tasks may have provided further cues that could have enabled the participants to become aware of the embedded grammatical feature at that point. Nevertheless, it is of interest that the group difference prevailed even on this last measure, and that the interview results correlated specifically with a critical morphological learning measure in the recognition task—namely, performance accuracy with items carrying a correct stem but an incorrect gender suffix. This suggests that a significant part of the participants' morphological awareness stemmed from the initial acquisition of the rule rather than on posttraining inference based on the previous learning tests. Finally, there might be an additional potential source of interference for the Spanish participants stemming from the partial inconsistencies of our versus the Spanish gender system. Our gender rule was fully consistent in the target items but the fillers were unmarked, while Spanish marks all nouns but includes a number of exceptions (e.g., although the word-final -o usually marks masculine gender, the word "mano", hand, is feminine) and exhibits a large variation of word-final segments. While we cannot rule out this possibility, we deem it as less likely given that both our artificial language and Spanish include irregularities, and because even our Finnish participants had been influenced by a language having a gender-marking system—namely, Swedish, which they all had studied at school. Again, new experiments with modified stimuli and other language pairs would be needed to address this issue.

In summary, we designed a new artificial morphological learning paradigm reminiscent of L2 word learning to explore possible cross-language differences in morphological acquisition in adults. The paradigm was successful in eliciting morphological learning in our participants, but with marked systematic differences between the Spanish and Finnish groups. The group differences suggest that for identification of an embedded suffix in a novel language, long-term general experience in morphological decomposition in native language can be more advantageous than familiarity with a similar grammatical feature in one's native tongue.

Original manuscript received 26 November 2013
Accepted revision received 1 August 2014

REFERENCES

Ambadiang, T. (1999). La flexión nominal. Género y número. In *Gramática descriptiva de la lengua española* (pp. 4843–4908). Madrid: Espasa Calpe.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321–324. doi:10.1111/1467-9280.00063

Comrie, B. (1989a). *Language universals and linguistic typology* (2nd ed.). Oxford: Basil Blackwell Ltd.

Comrie, B. (1989b). *The world's major languages*. London and New York: Routledge.

Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.

Endress, A. D., & Hauser, M. D. (2011). The influence of type and token frequency on the acquisition of affixation patterns: Implications for language processing. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*, *37*(1), 77–95. doi:10.1037/a0020210

Ferman, S., & Karni, A. (2010). No childhood advantage in the acquisition of skill in using an artificial language rule. *PLoS ONE*, *5*(10), e13648. doi:10.1371/journal.pone.0013648

Ferman, S., Olshtain, E., Schechtman, E., & Karni, A. (2009). The acquisition of a linguistic skill by adults: Procedural and declarative memory interact in the learning of an artificial morphological rule. *Journal of Neurolinguistics*, *22*(4), 384–412. doi:10.1016/j.jneuroling.2008.12.002

Hyönä, J., Laine, M., & Niemi, J. (1995). Effects of a word's morphological complexity on readers' eye fixation patterns. In J. M. Findlay, R. W. Kentridge, & R. Walker (Eds.), *Eye movement research: Mechanisms, processes and applications* (Vol. *6*, pp. 445–452). New York, NY: Elsevier.

Karlsson, F. (1983). *Finnish Grammar* (p. 222). Helsinki: WSOY.

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705–710. doi:10.3758/PBR.16.4.705

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review*, *19*(5), 935–941. doi:10.3758/s13423-012-0271-5

Laine, M. (1999). Meaning analysis of inflected words. *The Quarterly Journal of Experimental Psychology*, *52*(1), 253–259. doi:10.1080/713755799

Laine, M., Vainio, S., & Hyönä, J. (1999). Lexical access routes to nouns in a morphologically rich language. *Journal of Memory and Language*, *40*(1), 109–135.

Luraghi, S. (2011). The origin of the Proto-Indo-European gender system: Typological considerations. *Folia Linguistica*, *45*(2), 435–464. doi:10.1515/flin.2011.016

Marquis, A., & Shi, R. (2012). Initial morphological learning in preverbal infants. *Cognition*, *122*(1), 61–66. doi:10.1016/j.cognition.2011.07.004

Merkx, M., Rastle, K., & Davis, M. H. (2011). The acquisition of morphological knowledge investigated through artificial language learning. *The Quarterly Journal of Experimental Psychology*, *64*(6), 1200–1220. doi:10.1080/17470218.2010.538211

Mintz, T. H. (2013). The segmentation of sub-lexical morphemes in english-learning 15-month-olds. *Frontiers in Psychology*, *4*(24), 1–12. doi:10.3389/fpsyg.2013.00024

Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*(1), 271–280. doi:10.1016/j.cognition.2008.02.003

Pirkola, A. (2001). Morphological typology of languages for IR. *Journal of Documentation*, *57*(3), 330–348.

Rodríguez-Fornells, A., de Diego Balaguer, R., & Münte, T. F. (2006). Executive control in bilingual language processing. *Language Learning*, *56*(1), 133–190.

Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *22*(1), 1–29. doi:10.1191/0267658306sr259oa

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.

Soveri, A., Lehtonen, M., & Laine, M. (2007). Word frequency and morphological processing in Finnish revisited. *The Mental Lexicon*, *2*(3), 359–385.

Tamminen, J., Davis, M. H., Merkx, M., & Rastle, K. (2012). The role of memory consolidation in generalisation of new linguistic information. *Cognition*, *125*(1), 107–112. doi:10.1016/j.cognition.2012.06.014

Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio: The Psychological Corporation.

Whaley, L. J. (1997). *Introduction to typology: The unity and diversity of language*. Thousand Oaks, California: SAGE Publications.