Birdsong, D., & Flege, J. E. (2001). Regular-irregular dissociations in L2 acquisition of English morphology. *BUCLD 25: Proceedings of the 25th Annual Boston University Conference on Language Developmen*t (pp. 123-132). Boston, MA: Cascadilla Press.

Regular-Irregular Dissociations in L2 Acquisition of English Morphology

David Birdsong and James E. Flege
University of Texas-Austin and University of Alabama-Birmingham

1. Introduction

Under the dual-mechanism model of verb morphology, computation of regular verb past tense is a matter of rule-based, or symbolic, processing, and knowledge of irregular pasts involves access to individual lexical items that are stored in associative memory, and whose representation is sensitive to the frequency of the item. Within a single-system connectionist model, on the other hand, manipulation of symbolic compositional featues (stem + affix) is not posited for regulars: all past forms—regulars and irregulars alike—require retrieval from frequency-sensitive associative memory. For both models, the same assumptions are extended to the processing of regular versus irregular noun plurals.

Behavioral and neurofunctional evidence, most of it coming from L1 adults and children, suggests dissociations between rule-based and lexical knowledge, (e.g., Marcus et al., 1995; Pinker, 1999; Ullman et al., 1997). For the L2 context, behavioral evidence for the regular-irregular dissociation is found by Beck (1997), Marzilli & O'Brien (2000), and others. By and large, evidence of regular/irregular dissociations has been interpreted as consistent with the hybrid or dual-mechanism model of morphological representation (cf. Ellis & Schmidt, 1998; Elman et al., 1996).

The present investigation takes as its point of departure the Flege, Yeni-Komshian & Liu (1999) study, which looked at the end-state performance of Korean learners of L2 English. To date, this is the only published research that investigates the putative regular-irregular dissociation at the L2 end state as a function of age of arrival (AoA). Flege et al. (1999) broke down a 144-item subset of Johnson & Newport (1989)'s 276 stimuli into examples of Rule-based versus Lexical items. Rule-based items included regular *–ed* past verb forms and regular *–s* suffixation noun plurals. In the Lexical category were irregular verb pasts and noun plurals and lexical idiosyncrasies such as choice and placement of particles and prepositions (e.g., *jump over the fence* versus the ungrammatical *jump the fence over*; cf. the grammaticality of both *look over the car* and *look the car over*). Subjects rendered binary grammaticality judgments for randomly-ordered sentences, half of which were grammatical, half of which were ungrammatical counterparts. Presentation of stimuli was simultaneously oral and visual.

Accuracy figures (Flege et al., 1999, p. 91) for judgments of ungrammatical items reveal a clear dissociation in performance on regulars versus irregulars over AoA. For participants with AoA of approximately 7–8 years, performance on regulars and irregulars is identical (approximately 90% accuracy). However, as AoA increases, the Rule-based or regular items are more accurately judged overall; that is, they are less affected by increasing age of arrival than are the irregular items (Lexicals). Thus, at AoA=15 years, the regular items are judged at approximately 73% accuracy and the irregulars at 50% accuracy. At AoA=23 years, regulars are judged at approximately 70% accuracy, and performance on irregulars is below chance (approximately 48% accuracy).

We designed the present study primarily to try to replicate this observed dissociation in computation of regulars versus irregulars as a function of age of arrival. In particular, we wished to see if the Flege et al. (1999) results would reproduce for a set of items that were specifically designed for testing the regular/irregular distinction. In addition, we wished to revisit the question of whether L2 regularity effects are observed across linguistic categories (e.g., Marcus, 1995), in this instance across both noun plurals and verb pasts. Further, we were keenly interested in documenting frequency effects. By hypothesis, input frequency should be a factor in knowledge of irregular, but not regular forms. Given the ambivalent results reported by Beck (1997), however, the question remains unsettled for the L2 context. Finally, for these three types of hypothesized results, namely: the age by regular/irregular interaction, part of speech (linguistic category) effects, and input frequency effects, we wished to determine if results would generalize across different L1s. We therefore recruited two groups of participants, one whose native language was Spanish, and one whose native language was Korean.

2. Method

Our sample of educated Spanish Native Speakers (SNS; n=30) and Korean Native Speakers (KNS; n=30) at L2 asymptote was distributed along the biographical parameters given in Table 1.

Table 1. Design features: Participants' biographical parameters

	Age at Testing (Years)	Age of Arrival in U.S. (Years)	Length of Residence (Years)
10 KNS 10 SNS	16-26	6-10	10-16
10 KNS 10 SNS	21-31	11-15	10-16
10 KNS 10 SNS	26-36	16-20	10-16

Our first consideration in this design was avoiding the confound of Age of Arrival (AoA) with Length of Residence (LoR), which has been a drawback in some prior studies of age effects. We therefore held LoR constant, and within a narrow range of 6 years. We recruited participants from three non-overlapping AoA groups: 6–10 years (Child); 11–15 years (Adolescent), and 16–20 years (Post-Adolescent). From other studies we know that Age at time of Testing (AoT) is not strongly predictive of performance. However, AoT normally covaries with AoA, and so to unconfound somewhat these two factors, we stipulated that the AoT of the Child group should overlap with that of the Adolescents, which in turn overlaps with AoT of the Post-Adolescent group. Finally, by setting a minimum length of residence at ten years, we felt reasonably confident that participants were at asymptote in their L2 attainment (see Birdsong, 1999, ch. 1; Flege et al., 1999).

With the exception of two Spanish subjects and one Korean subject in the early-arriving group who had had one semester of college-level study, all participants had had at least one year of college in the U.S. (The youngest subject in our sample was 18-1/2 years of age.) We therefore felt that all subjects should have had sufficient U.S. education to expose them to the English vocabulary used in the low-frequency stimuli.

Stimuli included 80 multiple-choice items randomly presented on a Macintosh laptop computer using SuperLab software. Forty items exemplified English past tense morphology, with equal numbers of high- and low-frequency verbs. Within each frequency class, there were 10 regular and 10 irregular items. In addition, 40 items tested knowledge of English noun plurals; as with the verb past items, there was a counterbalanced distribution of high and low frequency, and of regular and irregular morphology. Sentence contexts were carefully controlled. For each verb item, an adverbial at the beginning of the sentence unambiguously indicated past reference. At the beginning of each noun item, a cardinal number or a quantifier such as "many" or "several" unambiguously indicated plurality. On all items, the five multiple choice possibilities were followed by a prepositional phrase or NP complement. Finally, all of the supplied contextualizations were constructed using high-frequency lexical items. (Twenty items exemplifying phrasal verbs and considered as a type of irregular construction were also included. We will not discuss these results here, except to say that performance on this class of items was comparable to that on the other Irregulars.) Two sample items are given in (1) and (2) below.

(1) Low Frequency Regular Noun Plural

There are five	a. knuckli	on each hand.
	b. knuckle	
	c. knuckles	
	d. knackle	
	e. knuckleses	

(2) High Frequency Irregular Verb Past

Yesterday the little girl	a. swim	for the first time.
	b. swam	
	c. swimmed	
	d. swims	
	e. swammed	

To determine an item's membership in high versus low frequency class, we consulted the Francis and Kucera frequency dictionary of English. High frequency items had a stem frequency up to ten times that of low frequency items. Participants were asked to respond as accurately and as rapidly as possible. Responses were indicated by pressing a, b, c, d, or e on a conventional computer keyboard. Participants were not permitted to go back and look at a previous item. There were five practice items. Our dependent measures were accuracy and response time.

3. Results

Table 2 displays a summary of the results of ANOVAs for accuracy data. Except where indicated, these results are observed for both the SNS and the KNS groups.

Table 2. Summary of ANOVAs for accuracy data: Significant main effects and interactions, *p* <.05

Factor	Significant effects
Age of Arrival	Child > Post-Adolescent (SNS); Child > Adolescent; Child > Post-Adolescent (KNS)
Regularity	Regulars > Irregulars
Frequency	High Frequency Items > Low Frequency Items
AoA X Regularity	Age Effect for Irregulars > Age Effect for Regulars
Frequency X Regularity	Irregular High Frequency > Irregular Low Frequency
Native Language Group X Part of Speech	Verbs > Nouns (KNS)
Native Language Group	KNS > SNS

Accuracy data and standard deviation figures for each of the above comparisons are given in Table 3.

Table 3. Accuracy, expressed as proportion correct, and standard deviations for relevant comparisons (see Table 2).

	SNS		KNS	
	Proportion Correct	Standard Deviation	Proportion Correct	Standard Deviation
Age of Arrival				
Child	.82	.25	.92	.13
Adolescent	.78	.29	.85	.19
Post-Adol	.71	.34	.81	.21

Regularity				
Regular	.92	.16	.95	.07
Irregular	.63	.33	.77	.21
Frequency				
High	.83	.24	.92	.12
Low	.72	.33	.80	.22
AoA X Regularity				
Child Reg	.94	.10	.98	.05
Child Irreg	.71	.29	.87	.15
Adol Reg	.89	.19	.95	.08
Adol Irreg	.68	.33	.74	.21
P-Adol Reg	.92	.17	.93	.07
P-Adol Irreg	.51	.34	.69	.23
Frequency X Regularity				
Hi Freq Reg	.94	.15	.97	.06
Hi Freq Irreg	.72	.27	.87	.13
Lo Freq Reg	.90	.17	.94	.07
Lo Freq Irreg	.54	.37	.66	.23
Native Language Group X Part of Speech				
Verb	.78	.28	.91	.14
Noun	.77	.31	.81	.21
Native Language Group				
	.77		.86	

As shown in Table 2, we observed main effects for Age of Arrival, Regularity, and Frequency. Recall, however, that we were interested primarily in interaction effects, especially a possible interaction of AoA and Regularity. For both the Korean Natives and the Spanish Natives, this interaction was significant. The interaction effect is consistent with the findings of Flege et al. (1999). There is an age effect over all items, but it is less pronounced for Regular items than for Irregulars.

We also wished to see if frequency effects would be observed across all items or would be restricted to Irregulars. Again, for both Korean and Spanish participants, we found a significant interaction. This result is consistent with the dual-mechanism model that predicts that computation of symbolically represented regulars will not be sensitive to stem frequency, whereas lexical representation in memory is dependent on input frequency.

Another result of interest is an interaction of Native Language Group by Part of Speech. Korean has a formal morphological mechanism for marking plurality on nouns, and it is regular. However, in most instances, plurals are not marked explicitly; rather, nouns retain singular morphology, and plurality is inferred from discourse cues. In contrast, verb past is explicitly marked in Korean. (In Spanish, both verb pasts and noun plurals are marked by morphological inflection.) Thus we wished to see if L1 influence would persist into the end state, in the form of Koreans' asymmetrical performance on nouns versus verbs. This difference was in fact observed.

Our accuracy data were supplemented by response latency (RT) data. As has been shown repeatedly in psycholinguistic research, response time can be a rather noisy variable, and our results are no exception. However, by and large, the RT results pattern in ways that resemb le the accuracy data. The main effects are identical across the two measures. For the interactions with RT data, not all effects are significant; however, they are all in the same direction as the accuracy data. In sum, the RT results are consistent with the accuracy results.

5. Additional analyses

As a secondary consideration, we wished to discover not only patterns of accuracy and RT, but also regularities in types of errors. It is important to gather data, for example, on whether verb past errors reflect a recurrent absence of overt past morphology, an over-generalization of *–ed* past, an inappropriate vowel alternation in the stem, and so forth. In short, we were hoping to collect data to supplement that which has already been analyzed in for early stages of L2A, which relates not only to computation of regulars and irregulars but to superordinate questions of end-state error patterns and of the separation of morphology and syntax.

Accordingly, for each item the distractor choices were varied systematically to represent inflecitonal permutations such as absence of marking, vowel alternation, overregularization, and so on. Examples of these permutations are given in Table 4.

Table 4. Permutations of response choices, by item type; correct responses in bold.

Item Type	Possible Responses	Example
Regular verb past	**• regular (-ed) suffix**	**helped**
e.g. help-helped	• vowel alternation	halp
	• no inflection	help
	• present tense inflection	helps
	• multiple past marking	helpeded
Irregular verb past (I)	• regular (-ed) suffix	swimmed

e.g. swim-swam	**• vowel alternation**	**swam**
	• no inflection	swim
	• present tense inflection	swims
	• multiple past marking	swammed
Irregular verb past (II)	• regular (-ed) suffix	cutted
e.g. cut-cut	• vowel alternation	cot
	• no inflection	**cut**
	• present tense inflection	cuts
	• multiple past marking	cotted
Regular noun plural	**• regular (-s) suffix**	**dogs**
e.g. dog-dogs	• vowel alternation	dig
	• no inflection	dog
	• irreglar suffix	dogen
	• multiple plural marking	dogses
Irregular noun plural(I)	• regular (-s) suffix	mouses
e.g. mouse-mice	**• vowel alternation**	**mice**
	• no inflection	mouse
	• irregular suffix	mousen
	• multiple plural marking	mices
Irregular noun plural(II)	• regular (-s) suffix	mooses
e.g. moose-moose	• vowel alternation	meese
	• no inflection	**moose**
	• irregular suffix	moosen
	• multiple plural marking	meeses
Irregular noun plural(III)	• regular (-s) suffix	phenomenons
e.g. phenomenon-	• vowel alternation	phenimenon
phenomena	• no inflection	phenomenon
	• irregular suffix	**phenomena**
	• multiple plural marking	phenomenas

Our results can be summarized briefly. For every class of item, and across both nouns and verbs, there is evidence of every type of error. However, remarkably, there is no distribution of errors that could be said to constitute a pattern. Error types are largely confined to individual items among low-frequency nouns and verbs, and are not observed across classes of items. Consider, for example, low frequency noun irregulars with final-vowel alternation:

Table 5. Examples of error types, low frequency nouns

Singular	ill-formed plural	(error type)
Crisis	crisis	(lack of inflection)
Criterion	criterions	(regular inflection)
Phenomenon	phenomenon	(lack of inflection)
	phenimenon	(vowel alternation)
	phenomenas	(multiple plural marking)
	phenomenons	(regular *–s* inflection)

5. Discussion

To conclude, let us return to the three interactions we were most interested in, each of which speaks to matters of interest in current acquisition theory. First, in the observed interaction of Frequency by Regularity, whereby only irregulars are sensitive to input frequency, we have unambiguous L2 evidence consistent with the hypothesized dissociation of regulars and irregulars laid out in the dual-mechanism model of linguistic computation. This evidence has not been reliably observed in previous L2 studies (e.g., Beck, 1997), perhaps because of procedural differences.

The interaction of Native Language Group X Part of Speech, whereby Koreans' performance on nouns was depressed relative to that on verbs, suggests straightforwardly an influence of native language. This finding adds to a growing body of experimental data that goes beyond the question of initial state L1 influence to demonstrate L1 influence at the end state. It also speaks to the need to address end-state study with sufficient granularity to capture this type of effect.

Finally, the Regularity by Age of Arrival interaction, whereby computation of irregulars is increasingly deficient over AoA, relative to computation of regulars, invites a bit of principled speculation. Cognitive neuroscience commonly distinguishes between declarative memory and procedural memory systems. According to Ullman et al. (1997), the declarative system underlies the learning and storage of arbitrarily-related information, such as names of people and novel facts, and can include morphological irregularities or grammatical idiosyncrasies. In the procedural system, learning and processing of motor, perceptual, and cognitive skills takes place. Since grammatical rules are like skills in that they require the coordination of procedures in real time, it is argued that regular noun and verb morphology is computed by the procedural memory system. Relating our results to Ullman et al. (1999)'s functional distinction, we have evidence that the declarative system may be more susceptible to aging effects than the procedural system.

The Regularity X AoA interaction likewise provides fodder for current L2A theory. Assuming that symbolic processes relating to grammar are handled by an innate Language Acquisition Device, the evidence presented here suggests that such a mechanism is relatively unaffected by age of acquisition, compared to the component that handles language-specific, irregular or idiosyncratic features. Thus, as some researchers posited a decade ago (e.g., Flynn & Manuel, 1991), it may be fruitful to investigate principled distinctions between different types of grammatical learning/representation/computation in L2A, especially in the context of critical period research (Birdsong, 1999; Birdsong & Molis, 2001).

References

Beck, Maria-Luise. (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence differences. *Second Language Research, 13,* 93-115.

Birdsong, David. (1999). *Second language acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Erlbaum.

Birdsong, David, & Molis, Michelle. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44,* 1-15.

Ellis, Nick C., & Schmidt, Richard. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes, 13,* 307-336.

Elman, Jeffrey L.; Bates, Elizabeth A.; Johnson, Mark H.; Karmiloff-Smith, Annette; Parisi, Domenico; & Plunkett, Kim. (1996). *Rethinking innateness*. Cambridge, MA: MIT Press.

Flege, James E.; Yeni-Komshian, Grace H.; & Liu, Serena. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41,* 78-104.

Flynn, Suzanne & Manuel, Sharon. (1991). Age-dependent effects in language acquisition: 'Critical period' hypotheses. In Eubank, Lynn (Ed.), *Point-counterpoint: Universal Grammar in the second language* (pp. 117-145). Amsterdam: Benjamins.

Johnson, Jacqueline S., & Newport, Elissa L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Marcus, Gary F. (1995). Children's overregularization of English plurals: A quantitative analysis. *Journal of Child Language, 22,* 447-459.

Marcus, Gary F.; Brinkmann, Ursula; Clahsen, Harald; Wiese, Richard; & Pinker, Steven. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology, 29,* 189-256.

Marzilli, Alisa, & O'Brien, Mary G. (2000). Plural morphology of German: A test of Marcus et al. (1995). Unpublished ms., University of Texas-Austin.

Pinker, Steven. (1999). *Words and rules.* New York: Basic Books.

Ullman, Michael T.; Corkin, Suzanne; Coppola, Marie; Hickok, Gregory; Growdon, John H.; Koroshetz, Walter J.; & Pinker, Steven. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*, 266-276.