Effects of Rule Instruction Versus Stimulus Highlighting: Are They Moderated by Time Pressure in Testing and Training?

Nora Presson and Brian MacWhinney

Carnegie Mellon University

1. Introduction

For adults learning a second language (L2), grammar is one of the most difficult skills to acquire, as reflected in non-nativelike grammar usage even in high-proficiency learners (e.g., Johnson & Newport, 1989; Abrahamsson & Hyltenstam, 2009). Some models attribute this difficulty to the low reliability, low salience, redundancy or optionality of some grammatical features (DeKeyser, 2005). Other models attribute adults' deficits in L2 grammatical processing to the attenuation after puberty of learning mechanisms that children use to acquire grammar (Johnson & Newport, 1989; Ullman, 2001). Whether or not adult grammar learning is fundamentally impaired, the relatively low level of aptitude in grammar processing makes instructional methods for improving grammatical representations an important problem in language education. The goal of the current study is to test whether explicit instruction of cues to grammatical category membership can improve performance in grammatical categorization, and whether adding a time pressure constraint during testing (and therefore encouraging more online processing) prevents learners from applying those explicitly learned cues, and whether time pressure during training improves learners' ability to adapt to that time pressure.

There is an abundance of evidence that making grammatical information explicit for learners can improve their use of that information (for meta-analyses see Norris & Ortega, 2000; Spada & Tomita, 2010). This is consistent with research that shows the importance of attention in uptake of relevant features (e.g., Schmidt, 1993; Taraban, 2003 for categories like gender in an artificial language).

In a prior experiment (Presson & MacWhinney, in preparation), we taught adults with no prior knowledge of French to categorize French nouns by grammatical gender on the basis of orthographic cues. Nouns in French are either masculine or feminine, and native speakers show predictable assignment of gender to nonce words (Tucker, Lambert, Rigault, & Segalowitz, 1968; Lyster, 2006). This is possible because the endings of nouns can serve as reliable predictors of gender, and these cues to gender can be described as phonological (Lyster, 2006) or orthographic (Holmes & Segui, 2004), and both formulations also reflect morphological information. In that study, explicit instruction combined with correctness feedback produced better learning and retention than did correctness feedback.

In the current study, we compare the effects of explicit instruction with the effects of cue highlighting (an input enhancement manipulation) to test whether explicit instruction adds value beyond directing attention to the target structure. Input enhancement (e.g., Sharwood Smith, 1993) is a common instructional manipulation that improves learning by directing attention to the correct part of the linguistic input.

One common critique of explicit instruction is that the resultant knowledge will be represented explicitly and that such representations will be difficult or impossible to use in online language processing (e.g., Krashen, 1981; 1994). Input enhancement, in contrast, does not require explicit instruction, so the resulting representations could be less declarative and therefore could be more

^{*}This project was supported by the Pittsburgh Science of Learning Center, which is funded by the National Science Foundation award number SBE-0836012. This project was also supported in part by a Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education (#R305B040063).

applicable in online language use. To test the ability to use grammatical category knowledge online, the current study uses time pressure during testing to ask whether that time pressure will reduce or eliminate any benefit of explicit instruction. Also, because performance under time pressure may be easier with practice, we compare both explicit instruction and input enhancement with and without time pressure during training.

We predict that explicit instruction will lead to more learning and greater retention than will an input enhancement using cue highlighting. In addition, we predict that time pressure during testing will not eliminate this advantage, and that time pressure during training will improve student ability to adapt to time pressure during testing.

2. Method

2.1. Participants

Participants were 95 (62 female, 33 male) students and community members at Carnegie Mellon University with no prior experience with French. Participants were recruited using online advertisements on a university message board. Exclusion criteria were any prior experience with French, any training in a Romance language (Italian, Spanish, or Portuguese), or having a native language that used a declensional system based on gender (e.g., German); participants whose native language was not English but did not use grammatical gender, such as Chinese or Japanese, were included. All participants were randomly assigned to one of four training conditions: Explicit Instruction + Time Pressure (n = 20), Explicit Instruction – Time Pressure (n = 21), Cue Highlighting + Time Pressure (n = 23), or Cue Highlighting – Time Pressure (n = 28).

2.2. Materials

The words presented in training and testing were representative of 24 orthographic cues to grammatical gender. These cues were chosen because of their high reliability (> 90% in an unpublished Lexique corpus analysis) and number of words containing the cue. For each cue, 14 words were presented in training, and 4 (different) words were presented during testing. Words were presented in French with no English translations. A list of all cues and words is included in Appendix A

2.3. Design

The experiment consisted of a pre-test, one training session, an immediate post-test, and a delayed post-test administered one week after training. Thus, the study had a 3 (pre-test / immediate post-test / one-week delayed post-test) x 2 (explicit instruction / cue highlighting) x 2 (with and without training time pressure) design. Each participant received post-tests both with and without time pressure but the pre-test was always administered without time pressure (time limit for each trial = 6000 ms). This was done to avoid frustration that was shown in pilot testing. Time pressure was operationalized as a 1400 ms response deadline, after which incorrect feedback was shown (in training) or the next trial was presented (in testing). This 1400 ms deadline was chosen because it was 2 SD above the mean post-test response latency for correct responses of 1050 ms measured in a previous study (Presson & MacWhinney, in prep).

2.4. Procedure

On Day 1, after giving informed consent, participants completed a pre-test with no feedback. Participants were told that French has masculine words, used with *le*, and feminine words, used with *la* (no examples were presented in instructions).

For all trials (both testing and training), participants were presented with a noun presented twice on the screen, once with *le* and once with *la*, and were asked "Which version is correct?" They were asked to press either the "M" key for *le* or the "F" key for *la*.

After completing the pre-test, participants trained for 42 minutes with feedback. Feedback consisted of green highlighting on the correct answer choice, and sound feedback depending on the accuracy of the correct response (a positive ding tone for correct and a negative buzz for incorrect). Participants who were randomly assigned to a training with time pressure condition had to complete each trial in 1400 ms or incorrect feedback was presented. For explicit cue instruction, incorrect feedback was accompanied by a statement of the correct orthographic cue (e.g., "-age \rightarrow le"). In contrast, the cue highlighting condition saw only correctness feedback after responding; however, the stimuli presented in the highlighting condition showed the relevant cue in capital letters (e.g., from AGE), and this property of the stimulus was always present, both before and after response.

2.5. Data Analysis

To test the effects of explicit cue instruction and time pressure during training, a mixed-effects model with a random intercept for each learner was used, with the probability of a correct response as a binomial dependent variable (meaning that all coefficients are logistic) or latency as a continuous dependent variable. For latency data, only correct responses were analyzed due to the inconsistency of latency measures for incorrect responses. Test time (pre-test / immediate post-test / delayed post-test), time pressure during learning, and explicit instruction versus cue highlighting were predictors.

Two separate models were estimated, for post-tests with time pressure and post-tests without time pressure. In both cases, pre-test data were used as baseline, even though the pre-test was only given with no time pressure.

3. Results

3.1. Descriptive statistics

		Pretest		Immediate Post-Test				One-Week Delay			
		Untimed		Untimed		Timed		Untimed		Timed	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Overall (N=95)	Accuracy	.53	.05	.82	.04	.77	.08	.77	.04	.74	.05
	Latency*	1855	122	1195	66	860	27	1169	68	873	25
Explicit +	Accuracy	.52	.11	.85	.08	.82	.08	.78	.09	.77	.09
Timed (N=20)	Latency	2002	260	1084	111	875	217	1095	129	890	229
Explicit -	Accuracy	.53	.11	.86	.08	.78	.09	.81	.09	.77	.09
Timed (N=21)	Latency	1960	279	1175	150	846	294	1215	145	869	234
Highlighting +	Accuracy	.55	.10	.81	.08	.76	.09	.74	.09	.72	.09
Timed (N=23)	Latency	1827	247	1194	130	843	289	1174	146	867	284
Highlighting -	Accuracy	.52	.09	.78	.08	.74	.08	.74	.08	.70	.09
Timed (N=28)	Latency	1686	208	1306	132	872	260	1190	129	867	242

^{*}Latency (ms) analyzed for correct responses only

Overall mean proportion correct at pre-test was .53 (SE = .05), indicating that participants were not responding above chance to the overall word set, and therefore could be responding randomly, though there was variability in accuracy by gender cue that could be considered for future analysis of pre-existing biases. There were no pre-test differences among the instructional conditions (largest t(49) = 0.21, p = .83). Note that the overall mean latency on untimed tests after training was shorter than the response deadline (1400ms).

3.2. Tests without time pressure

3.2.1. Accuracy

The untimed tests showed significant learning, as indicated by the significant coefficient for both

¹ For latency (i.e., non-binomial outcome) models, all p-values were estimated using Markov Chain Monte Carlo simulation with the languageR package (http://cran.r-project.org/web/packages/languageR/index.html) in R.

post-tests compared to pre-test (for immediate post-test β = 1.23, p < .0001). Evidence for retention comes from the significant coefficient of delayed post-test compared to pre-test baseline (β = 0.96, p < .0001). There was no main effect of either explicit instruction (β = 0.07, p = .65) or time pressure during training (β = 0.12, p = .43). There was a significant interaction of test time (immediate post-test and one-week delayed post-test) and explicit instruction, reflecting that the conditions were equivalent at pre-test, but that the explicit instruction conditions showed greater improvement at immediate post-test than did the cue highlighting conditions (β = 0.60, p < .0001). The advantage of explicit instruction was also present at the one-week delayed post-test (β = 0.39, p < .0001). There was also a marginally significant three-way interaction among pre-test to immediate post-test improvement, explicit instruction, and time pressure during training (β = -0.25, p = .085), reflecting the trend that the explicit instruction / no time pressure during training group showed the highest average accuracy at immediate post-test. Figures 1a and 1b show this pattern for the tests without time pressure and with time pressure, respectively.

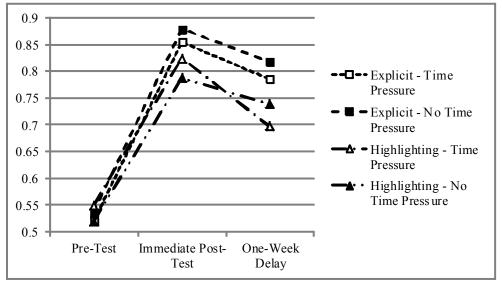


Figure 1a. Explicit instruction shows greater improvement in accuracy from pre- to post-test regardless of time pressure during training (for untimed tests).

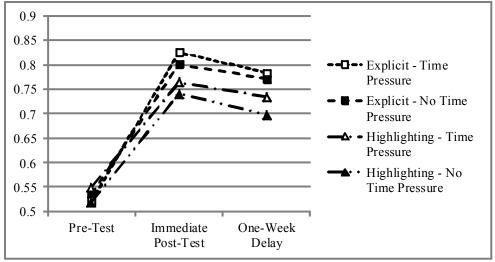


Figure 1b. Explicit instruction shows greater improvement from pre- to post-test regardless of time pressure during training (for timed tests).

3.2.2. Latency of correct responses in untimed tests

Analysis of latency data show a significant speed-up in the untimed tests after training, as indicated by a significant coefficient for immediate post-test (β = -403.4, p < .0001). This speed-up was maintained at the one-week delayed post-test (β = -487.7, p < .0001). There was a main effect of explicit instruction (β = 272.2, p = .012) that must be interpreted in light of interactions between each post-test and explicit instruction (immediate post-test β = -411.9, p < .0001; one-week delayed post-test β = -348.7, p < .0001). After training, the decrease in latency was larger for explicit instruction groups than for cue highlighting.

There was no main effect of time pressure during training (β = 138.7, p = .21), but there was a significant interaction of time pressure during training and change in latency from pre-test to each of the two post-tests (immediate post-test β = -263.0, p < .0001; delayed post-test β = -163.9, p < .001), such that the groups that trained under time pressure responded faster on the untimed post-tests than those who trained without time pressure.

There was a significant three-way interaction among immediate post-test, explicit instruction, and time pressure during training ($\beta = 152.6$, p = .007). At the immediate post-test, the lowest response latencies were for the group who trained with explicit instruction and no time pressure. This mirrored the finding for accuracy reported earlier. However, the latency advantage disappeared at the delayed post-test ($\beta = 83.5$, p = .16).

3.3. Tests with time pressure

3.3.1. Accuracy

The analysis of post-tests with time pressure, also showed learning from pre-test to immediate post-test ($\beta = 0.98$, p < .0001). This learning was maintained to the one-week delayed post-test ($\beta = 0.75$, p < .0001).

There was no main effect of either explicit instruction ($\beta = 0.07$, p = .61) or time pressure during training ($\beta = 0.12$, p = .36). However, there was a significant interaction between explicit instruction and improvement from pre-test to immediate post-test ($\beta = 0.27$, p = .004), such that improvement from pre-test on the timed tests was larger for explicit instruction than for cue highlighting conditions. This interaction was also present at the delayed post-test ($\beta = 0.31$, p = .002).

Contrary to our predictions, there was no effect of time pressure during training on performance on the time pressure post-tests (immediate post-test $\beta = 0.001$, p = .99; delayed post-test $\beta = 0.06$, p = .53). There were also no three-way interactions (immediate post-test $\beta = 0.22$, p = .11; delayed post-test $\beta = 0.07$, p = .62).

3.3.2. Latency of correct responses

The latency model for the post-tests with time pressure showed a large speed-up between pre-test and both immediate post-test (β = -807.6, p < .0001) and delayed post-test (β = -835.8, p < .0001). This reflects the addition of the time pressure constraint, so the amount of speed-up cannot be fully attributed to learning; however, the coefficient for delayed post-test is more negative than that for immediate post-test, suggesting that participants also sped up at the delayed post-test.

There was a significant main effect of explicit instruction (β = 281.2, p < .0001) such that explicit instruction took longer. However, this must be interpreted in light of a significant interaction between post-tests and explicit instruction (immediate post-test β = -348.0, p < .0001; delayed post-test β = -375.4, p < .0001), reflecting the fact that explicit instruction led to more speed-up (or better adaptation to the time pressure constraint during testing) than did input enhancement with cue highlighting.

There was no main effect of time pressure during training (β = 141.6, p =.11), but there was a significant interaction of time pressure during training with post-test occasions (immediate post-test β = -202.4, p < .0001; delayed post-test β = -111.4, p = .0015), showing that participants who trained with time pressure showed greater speed-up (or better adaptation to the time pressure constraint during testing) than did participants who trained with no time pressure.

The two-way interaction between explicit instruction and training time pressure was not significant (β =-94.2, p =.46). However, there were significant three-way interactions among explicit

instruction, time pressure during training, and each of the two post-tests (immediate post-test β = 215.2, p < .0001; delayed post-test β = 184.1, p = .0003), reflecting that the largest speed-up (or adaptation to the time pressure constraint during testing) was for the group who trained with explicit instruction and no time pressure.

4. Discussion

The current study compared the effects of explicit instruction with gender cues to the use of input enhancement with capitalization, and found that explicit instruction led to higher accuracy for both tests with and without time pressure. Both conditions led to successful learning, confirming that directing learners' attention to the relevant feature can lead to uptake of those features with correctness feedback alone; however, explicit instruction with those cues led to even greater learning, and greater retention one week after training. This result suggests that the representations formed with explicit instruction are stronger than those formed with cue highlighting. Moreover, the addition of a time pressure constraint during testing did not eliminate the advantage of explicit instruction. This contradicts the prediction in Krashen (1981; 1994) that explicit instruction will be difficult to use in online language processing because the time demands do not allow the use of declarative knowledge.

The findings for response accuracy were mirrored for response latency. That is, explicit instruction also led to faster responses than did stimulus highlighting. Taken together, the latency and accuracy results demonstrate that when explicit instruction led to better categorization accuracy, this was not accompanied by either a slowdown in performance or impairment in dealing with the added time pressure constraint in the post-tests. In fact, the latency models suggest that explicit instruction led to faster performance than cue highlighting both with and without time pressure in testing. Thus, there was no evidence that explicit instruction resulted in a speed-accuracy tradeoff and no evidence that explicit instruction led to representations that were not usable in online language processing.

In future studies, several features of this design could be changed to reflect more naturalistic properties of the grammatical system. First, the current study did not present any exceptions to the gender cues, such that all cues seemed to have 100% reliability. In real language, of course, this is not the case, and it is reasonable to ask whether exceptions would harm the application of explicitly represented grammatical knowledge more than less explicit representations such as those produced by input enhancement.

Second, words in the current study were presented with their definite articles, but in the absence of a larger sentence context (because the population had no prior French knowledge). The use of explicitly learned cues could be more difficult in a language context. However, the lack of response latency penalty for explicit instruction makes this possibility less likely.

In sum, the current study supports the idea that explicit instruction adds value to student learning beyond redirecting their attention to relevant target features. By presenting a direct decision rule, students may be able to form a more strongly represented schema. Explicit instruction of grammar is already pervasive in language education, and this study suggests that such instruction can improve student learning.

5. Appendix A: Cues and words presented

Cue	Training Words	Test Words
1age (M)	garage, étage, fromage, paysage, recyclage, nuage, visage, potage, avantage, outrage, maquillage, ménage, repassage, stage	apprentissage, tapage, clivage, otage
2ance (F)	avance, confiance, outrance, aisance, béance, résistance, enfance, naissance, espérance, plaisance, puissance, resonance, souffrance, connaissance	protuberance, abondance, croyance, tendance

3d (M)	canard, billard, plumard, bord, traînard, nid, lard, retard, papelard, renard, froid, chaud, bled, accord	nid, dard, bagnard, gond
4é (M)	marché, congé, procédé, préjugé, peché, bébé, canapé, café, cliché, résumé, scellé, blé, defilé, ciné	triplé, curé, envoyé, thé
5ée (F)	poupée, fée, armée, nausée, allée, chevauchée, coulée, vallée, invitée, fumée, musée, année, idée, journée	mosquée, armée, musée, buée
6ère (F)	panthère, prière, colère, banquière, bière, verrière, lisière, infirmière, étagère, matière, bannière, ferrière, misère, rivière	croisière, carrière, frontière, paupière
7en / ien (M)	musicien, pharmacien, informaticien, chien, technicien, généticien, mathématicien, politicien, doyen, magicien, statisticien, moyen, terrien, gardien	citoyen, rien, chirurgien, soutien
8eur (M)	auteur, coiffeur, acteur, coeur, tricheur, ordinateur, traducteur, fugueur, dîneur, présentateur, sauteur, ingènieur, tailleur, joueur	distributeur, transbordeur, entrepreneur, choeur
9i / oi (M)	ami, cari, convoi, parti, abri, défi, céleri, roi, emploi, quai, pari, ennemi, essai, vrai	canari, apprenti, alcali, frai
10ie (F)	hystérie, analogie, confiserie, académie, monnaie, pharmacie, allergie, bougie, garderie, hypocrisie, anarchie, sortie, envie, infamie	mercerie, jacasserie, agonie, apologie
11ier (M)	officier, grenier, quartier, caissier, banquier, tablier, collier, gravier, papier, barbier, cavalier, fournier, chemisier, escalier	terrier, fumier, lunetier, oreiller
12in / ain (M)	pain, raisin, terrain, foin, gradin, fin, destin, étain, ravin, brin, chagrin, jardin, requin, bain	parrain, tremplin, parchemin, puritain
13ine / aine (F)	cuisine, migraine, usine, médecine, rétine, épine, gazoline, rétine, vitrine, porcelaine, bassine, piscine, farine, poitrine	échine, crépine, voisine, isnuline
14isme (M)	capitalisme, judaïsme, cyclisme, criticisme, égotisme, prisme, charisme, futurisme, activisme, cubisme, ironisme, négativisme, autisme, baptisme	cosmopolitisme, mutisme, journalisme, élitisme
15me (M)	problème, rhume, drame, calme, dilemme, drame, poème, sarcasme, fantôme, rythme, charme, thème, terme	barème, diplôme, asthme, dome

16nne (F)	couronne, canne, ancienne, obsidienne, consonne, antenne, colonne, électricienne, musicienne, paysanne, magicienne, piétonne, lionne, nonne	tonne, practicienne, collégienne, espionne
17on (M)	pardon, faucon, sermon, poumon, avion, cordon, citron, dicton, manchon, savon, démon, pantalon, salon, menton	capuchon, goudron, torchon, charbon
18se (F)	défense, course, réponse, bosse, graisse, crise, danse, bourse, thèse, parenthèse, cuisse, ruse, anse, transe	entorse, chausse, pause, crevasse
19esse (F)	princesse, vitesse, politesse, vieillesse, hôtesse, déesse, presse, messe, promesse, maîtresse, adresse, justesse, finesse, sagesse	prêtresse, grossesse, faiblesse, comtesse
20t (M)	pont, bruit, trajet, droit, berlingot, commissariat, croyant, carnet, billet, point, mandat, syndicat, passeport	achat, doigt, mot, saut
21té (F)	beauté, santé, parité, habilité, densité, liberté, identité, humanité, proximité, impureté, agilité, publicité, egalité, cite	avidité, societé, rivalité, comptabilité
22tion / sion (F)	condition, pression, pension, alimentation, discrimination, préoccupation, immigration, salutation, désorientation, signification, organization, operation, détection, audition	affirmation, attention, denegation, inundation
23tre (M)	ministre, centre, prêtre, pitre, ârtre, antre, arbiter, monstre, ventre, orchestre, filter, astre, pupitre, désastre	cintre, peintre, theater, meurtre
24u (M)	inconnu, tissu, trou, élu, tonneau, alu, faisceau, panneau, préau, cru, radeau, farfelu, parvenu, anneau	essieu, flou, taureau, noyau

References

- Abrahamsson, Niclas, & Hyltenstam, Kenneth. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249-306.
- DeKeyser, Robert M. (2005). What makes learning second-language grammar difficult? A review of issues. Language Learning, 55, 1-25.
- Holmes, Virginia M., & Segui, Juan. (2004). Sublexical and lexical influences on gender assignment in French. *Journal of Psycholinguistic Research*, 33, 425-457.
- 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.
- Krashen, Stephen. (1981). Second Language Acquisition and Second Language Learning. Oxford: Pergamon Press
- Krashen, Stephen. (1994). The Input Hypothesis and its rivals. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-78). San Diego: Academic.

- Lyster, Roy. (2004). Predictability in French gender attribution: A corpus analysis. *French Language Studies*, 16, 69-92.
- Norris, John M., & Ortega, Lourdes. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.
- Presson, Nora, & MacWhinney, Brian (in preparation). Learning grammatical gender: The effects of rule instruction.
- Sharwood Smith, Michael (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165-179.
- Spada, Nina, & Tomita, Yasuyo. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, 263-308.
- Schmidt, Richard. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- Taraban, Roman. (2004). Drawing learners' attention to syntactic context aids gender-like category induction. Journal of Memory and Language, 51, 202-216.
- Tucker, G. Richard., Lambert, Wallace E., Rigault, Andre, & Segalowitz, Norman (1968). A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, 7, 312-316.
- Ullman, Michael T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30, 37-69.

Selected Proceedings of the 2010 Second Language Research Forum: Reconsidering SLA Research, Dimensions, and Directions

edited by Gisela Granena, Joel Koeth, Sunyoung Lee-Ellis, Anna Lukyanchenko, Goretti Prieto Botana, and Elizabeth Rhoades

Cascadilla Proceedings Project Somerville, MA 2011

Copyright information

Selected Proceedings of the 2010 Second Language Research Forum: Reconsidering SLA Research, Dimensions, and Directions © 2011 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-448-5 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper. Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press. To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Presson, Nora and Brian MacWhinney. 2011. Effects of Rule Instruction Versus Stimulus Highlighting: Are They Moderated by Time Pressure in Testing and Training? In *Selected Proceedings of the 2010 Second Language Research Forum*, ed. Gisela Granena et al., 105-113. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2619.