

# “*The Rain in Spain*” Updated? An Elocution Drill for Efficiently Teaching English Consonant Clusters

Omar A. Alkhonini<sup>1</sup>, Douglas J. Wulf<sup>2</sup>

<sup>1</sup>Majmaah University, Al Majma'ah, KSA

<sup>2</sup>Linguistics Program, Department of English, George Mason University, Fairfax, VA, USA

Email: om.alkhonini@mu.edu.sa, oalkhoni@gmu.edu, dwulf@gmu.edu

**How to cite this paper:** Alkhonini, O. A., & Wulf, D. J. (2018). “*The Rain in Spain*” Updated? An Elocution Drill for Efficiently Teaching English Consonant Clusters. *Open Journal of Modern Linguistics*, 8, 231-241. <https://doi.org/10.4236/ojml.2018.86020>

**Received:** November 8, 2018

**Accepted:** December 18, 2018

**Published:** December 21, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

This small-scale longitudinal study investigates a technique for assisting English learners to achieve more native-like pronunciation of onset consonant clusters. Study participants were ESL students with L1s that lack onset clusters (Arabic and Mandarin Chinese). In the study’s instruction phase, participants followed the venerable pedagogy of reciting an elocution drill, a 10-word sentence containing one occurrence each of the five three-consonant onsets of English, but containing none of the two-consonant onsets. This pedagogical technique conforms to the concept of language acquisition under Skill Acquisition Theory (SAT). Furthermore, Universal Grammar (UG) and the Markedness Hypothesis predict that by practicing only the more marked three-consonant onsets, two-consonant onsets might also be acquired despite not practicing them. Indeed, the results show statistically significant modification toward more native-like pronunciation of both three-consonant and two-consonant onsets from the pre-test to the immediate post-test and the delayed post-test, given 24 weeks later.

## Keywords

Second Language Acquisition, Learning Strategies, Skill Acquisition Theory, Markedness Theory, Universal Grammar, Language Pedagogy, English Onset Consonant Clusters, English as a Second Language

## 1. Introduction

In the 1964 film *My Fair Lady*, Henry Higgins, a phonetics professor, puts the unfortunate Eliza Doolittle through an intensive regime of elocution drills, the most memorable of which is “*The rain in Spain stays mainly in the plain*”. His

aim was to convert Eliza's Cockney pronunciation, which uses [aɪ] in words like "rain" and "Spain", into the socially prestigious Received Pronunciation, which uses [eɪ] instead. The film portrays the drills as something of a torture, especially as Eliza is made to endure hour upon hour of them. Despite this negative depiction, we suggest that a modern (and much gentler) update of this technique, informed by theories of cognitive science and linguistics, might serve a useful role in second-language (L2) pronunciation instruction.

Today, we rarely seek to bring about a massive transformation of a student's pronunciation, as in Eliza's case. However, even with the modest aim of fostering greater comprehensibility, we still attempt to encourage learners to achieve more native-like production. Non-native pronunciation does not necessarily lead to miscommunication as the speaker's intended meaning is often recoverable from context, but the listener must often work harder to recover the intended meaning. For instance, if "steam" instead of "stream" is produced in the utterance "She had tears steaming down her face", the intended meaning is recoverable in spite of the simplification of the onset consonant cluster, but only with a greater processing effort by the listener.

Interest in pronunciation instruction seems to be on the rise. Based on a survey of ESL teachers about teaching pronunciation, Foote, Holtby, and Derwing (2012) advocate that more attention be given to pronunciation instruction, that this instruction be better and more consistently integrated into classes, that pronunciation be taught by teachers trained in pronunciation instruction, and that teachers provide explicit feedback while prioritizing pronunciation issues for students.

Conspicuous hurdles in teaching pronunciation are sounds and sound combinations that linguists identify as *phonologically marked*. These are elements of the phonology that are less commonly found in languages generally and that are regarded as objectively more difficult to acquire. For example, whereas learners rarely have trouble with the unmarked alveolar fricative [s], as in *sink*, they tend to have considerably more difficulty with the marked interdental fricative [θ], as in *think*. Likewise, consonant clusters are often a challenge. English employs a large assortment of clusters of two and three consonants, many of which are disallowed by the phonological systems of other languages, such as Arabic and Mandarin.

This paper reports the results of a small-scale longitudinal study investigating an efficient technique for assisting English learners to achieve more native-like pronunciation of consonant clusters in onset position. Like the famous "Rain in Spain" drill, study participants practiced a 10-word sentence containing one occurrence each of the five three-consonant onset clusters of English: "A *scrod* will *splash*; a *squid* will *spray* a *stream*". However, unlike Henry Higgins' pedagogy, participants only practiced this drill for a fraction of a single half-hour instruction session.

Despite encountering this drill only once and to a minimal extent, results of both an immediate post-test as well as a delayed post-test, given 24 weeks later,

indicate a statistically significant modification toward more native-like pronunciation of the three-consonant onsets practiced. More remarkably, both post-tests also show similar modification of two-consonant onsets that were tested but not practiced at all via this drill. This intriguing result conforms to predictions of Universal Grammar (Chomsky, 1965) and the Markedness Hypothesis (Greenberg, 1976) in that acquisition of more marked forms potentially entails automatic acquisition of less marked forms also.

Brief use of this particular drill thus could be a highly efficient means of teaching onset clusters. Also, according to Foote et al. (2012), the most difficult problem in teaching pronunciation is determining what issues to emphasize when students have different L1s and cultural backgrounds. Given that both L1 Arabic and L1 Mandarin participants demonstrated statistically significant modification, this drill seems potentially helpful for students with widely differing L1s and cultures.

This technique is also in conformity with how language acquisition is conceptualized under Skill Acquisition Theory (SAT) (Anderson, 1976; DeKeyser, 1998; McLaughlin, 1987). SAT, a theory from cognitive science, treats language acquisition as just one example of the more general phenomenon of skill acquisition. In this process, learners first learn consciously what a skill requires of them (declarative knowledge). Through practice, learners then attempt to convert this into skill performance (procedural knowledge). Eventually, after extensive practice, the skill may reach automatization, such that it can be executed subconsciously.

For example, in learning to play the piano, one must first understand what to do (declaratively), then attempt to use this in initial attempts to play (procedurally), and finally, after extensive practice, develop smooth performance under subconscious control (automatically). Similarly, via an elocution drill, a student could move from declarative knowledge about consonant clusters, to their proceduralization in pronunciation, and eventually to an ability to produce them automatically in normal speech.

## 2. The Experiment

### 2.1. Method

All phases of the experiment took only 70 minutes total. There were four phases: pre-test (12 minutes), instruction (30 minutes), immediate post-test (10 minutes), and delayed post-test (18 minutes). The study was conducted in a George Mason University classroom. One of the authors conducted the study. The tests were presented to participants in paper form. For each test, participants were asked to read a list of words slowly with a brief pause between each item. A MacBook Air computer-based microphone was used to record the data and Praat software was used (Version 6.0.17; Boersma & Weenink, 2016).

### 2.2. Participants

The number of participants in this study was small. However, the aim was not to

provide comprehensive data on the extent that pronunciation might be modified. Rather, the objective was merely to locate any positive indication that this technique might be valuable. Might such a drill/mnemonic be helpful for any learner? If a positive indication were found, more extensive research might be undertaken.

There were 13 participants, consisting of 10 L1 Arabic speakers (two female, eight male) and three L1 Mandarin speakers (one female, two male), enrolled in the same beginner level ESL class at George Mason University. A placement test the university administered had assigned these students to this class. For all participants, the highest level of education completed was high school, except one who had completed some college and one with a bachelor's degree. Their ages were between 18 and 35 years old, but 10 participants were between 18 and 25. They had all started learning English between the ages of 10 to 15 and reported having no known hearing or speech problems.

For the delayed post-test, only 10 participants were available to be tested. The ESL instructor of these participants was asked if they had received any instruction in the previous 24 weeks targeting the pronunciation of onset clusters, and the instructor reported they had not. Incidentally, after they had performed the delayed post-test, several participants remarked that they had never been informed that their production of onsets differed from what native speakers of English produce.

### **2.3. Instrument**

The stimuli consisted of 11 items in the pre-test and 11 in the immediate post-test. Four items in each test were used as fillers and were not a target of the study. Of the seven target items, two contained two-consonant onsets, and five contained three-consonant onsets. The purpose of testing two-consonant onsets was to determine whether practicing three-consonant onsets might automatically improve pronunciation of two-consonant onsets also, as suggested by UG and the Markedness Hypothesis.

Two word lists, Form A and Form B (see [Appendix A](#)), were used for both pre-test and immediate post-test in the following way: seven participants were given Form A as the pre-test and Form B as the immediate post-test, and the other six had Form B as pre-test and Form A as immediate post-test. This was to verify that items in both lists represented approximately the same level of pronunciation difficulty for participants. For the delayed post-test, participants read both Forms A and B.

### **2.4. Procedure**

Participants provided their demographics and were told they would be audio-recorded three times: once before the instruction and twice after it. Participants waited outside the classroom and were called in one at a time to be recorded reading either Form A or Form B aloud as the pre-test. Each participant

was recorded separately to avoid having pronunciation potentially influenced by others.

Next was the instruction phase. This included explicit instruction about how to accurately produce English consonant clusters. Non-native productions were described and compared with native-like ones. Specifically, example words from neither form were written on a whiteboard, and participants were asked to read them aloud. The researcher also read them aloud and explained the differences between non-native and native-like pronunciation. Non-native production is often characterized by the insertion of a vowel into the cluster either before (i.e., VCCC) or inside (e.g., CVCC) the cluster.

It was also noted how non-native-like production could potentially be misunderstood by native speakers. For instance, adding a vowel inside the onset cluster in the word “square” [skwɛɪ] might cause this to be perceived as “sick where” [sikwɛɪ]. Admittedly, context usually helps listeners interpret what speakers want to communicate, but producing the cluster without this vowel insertion reduces the interpretative burden placed on the listener.

To this point, the instruction was just straightforward explanation/demonstration. However, in conformity with SAT, it is necessary to package relevant information in such a way as to facilitate its proceduralization and automatization. Therefore, the elocution drill was taught to serve both as a mnemonic device as well as a vehicle for proceduralizing the three-consonant onsets. The drill (“*A scrod will splash. A squid will spray a stream*”) contains one occurrence each of the five three-consonant onsets of English. By design, the sentence contains only the three-consonant onsets of English with none of the numerous two-consonant onsets of English.

Participants were asked to read this sentence carefully, aiming for native-like pronunciation. They practiced it for a few minutes and tried to memorize it. The researcher then read it aloud three times, and participants were asked to read it aloud. The researcher checked the pronunciation of each participant to verify that it was native-like. The researcher also answered any relevant questions raised. Then, the immediate post-test was administered. Participants were again recorded individually. As noted, participants who had read Form A for the pre-test, read Form B as the immediate post-test, and vice versa.

Twenty-four weeks later, 10 participants recorded a delayed post-test by reading both Form A and Form B in the same order that he or she had encountered them previously. Of these participants, just one was an L1 Mandarin speaker, and nine were L1 Arabic speakers.

The study was only concerned with the production of onsets. Thus, if a participant exhibited pronunciation that was not native-like in the coda or the peak of the syllable (e.g., changing vowel quality), provided the onset cluster was pronounced in native-like fashion, this was counted as native-like. Two participants in the immediate post-test repeated the words “*split*” and “*splint*”, such that the first production was not native-like, but the second was. Because the

study was investigating participant awareness of pronunciation, and self-correction is an indicator of such awareness, the production was counted as native-liked based on the second (corrected) production.

## 2.5. Data

Three types of non-native-like patterns (i.e., learner strategies or cluster simplification tools) were found in the data: epenthesis, deletion, and feature change. The most common was epenthesis, occurring 44 times (78%); followed by deletion, occurring 11 times (20%); and, feature change, occurring only once (2%). Among others, Weinberger (1987) and La Cruz and Savaria (2010) indeed claim that when speakers encounter clusters, the most likely modifications are epenthesis (insertion of a sound), deletion (the omission of one or more sounds), and feature change.

Epenthesis occurred almost always either before or after the first consonant of the cluster (e.g., VCCC or CVCC), though there were two instances of vowel insertion after the second consonant (i.e., CCVC). Both instances were with the word “*stream*” [stæim], and occurred only in the pre-test by two L1 Arabic participants. Deletion occurred with several participants in multiple words. The word exhibiting the most deletions was “*squad*”, which was simplified by six participants in the pre-test by deleting the third consonant [w]. Feature change was observed only in the pre-test with one participant pronouncing “*flag*” by darkening the /l/ to [ɫ].

In the pre-test and immediate post-test, a total of 182 tokens were produced by the 13 participants (i.e., seven in the pre-test and seven in the immediate post-test from each participant). In the delayed post-test, a total of 140 tokens were produced by the 10 participants (i.e., again, 14 from each participant). As noted, the delayed post-test included the words from both the previous tests. Thus, in comparing performance, it is clearer to examine percentages of native-like productions rather than their quantity. Table 1 displays percentages of native-like productions in each test. Thus, these were productions that did not exhibit epenthesis, deletion, or feature change.

Investigating the results further, a repeated-measures analysis variance was performed. Mauchly’s Test of Sphericity indicate that the assumption of sphericity had been violated,  $p < .001$ . Therefore, a Greenhouse-Geisser correction was used. This indicated significant mean differences in the three tests:  $F(1.76, 15.82) = 36.53$ ,  $p < .001$ . Because three participants did not participate in the delayed post-test, they were omitted from this analysis (see Table 2). A post hoc test using Bonferroni correction revealed an increase in the mean of native-like productions between the pretest and the immediate post-test ( $M = 47.1$ ,  $SD = 5.98$  for the pretest vs  $M = 94.3$ ,  $SD = 3.18$  for the immediate post-test), and this difference was statistically significant ( $p < .001$ ). The delayed post-test showed a slight decrease of its mean ( $M = 81.5$ ,  $SD = 4.94$ ) from the immediate post-test, but this was not significantly different ( $p > .05$ ). Yet, the results of the delayed post-test still retained a significant difference from the pre-test ( $p < .01$ ).

**Table 1.** Percentages of native-like productions per test for each participant.

Participant	Pre-Test	Imm. Post-Test	Del. Post-Test
1	29%	71%	71%
2	14%	100%	93%
3	71%	100%	-
4	43%	86%	93%
5	57%	100%	100%
6	71%	100%	86%
7	71%	100%	86%
8	57%	100%	93%
9	57%	100%	79%
10	29%	86%	64%
11	14%	57%	-
12	43%	100%	50%
13	43%	100%	-
Total	46%	92%	81%

**Table 2.** Descriptive statistics.

	Mean	Std. Deviation	N
Pre-Test	47.10%	18.91766%	10
Imm. Post-Test	94.30%	10.04490%	10
Del. Post-Test	81.50%	15.61516%	10

Due to the small sample size, no detailed conclusions can be drawn about the relative effectiveness of this technique for various groups tested in this study. That said, as **Table 1** indicates, without exception, each participant produced a higher percentage of native-like productions in the immediate post-test than in the pre-test. Whereas no participant achieved native-like production with all pre-test items, nine participants managed this in the immediate post-test. Although percentages of native-like productions declined in the delayed post-test for all but two participants, in no instance did this percentage fall to a level at or below that participant's pre-test percentage. We thus have a preliminary finding that this technique could be effective despite differences such as L1, age, and/or gender.

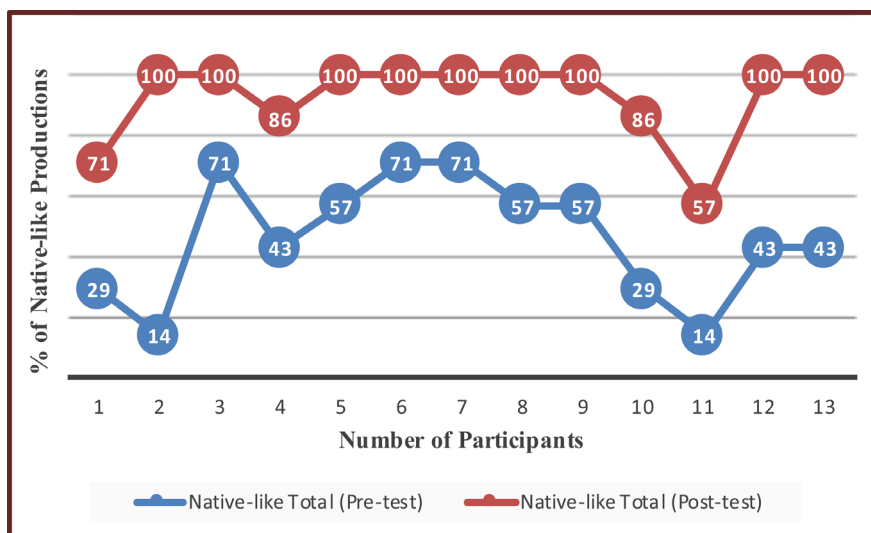
Finally, as **Table 3** indicates, there were no large differences between participants who took the tests in AB order in comparison with those who took the tests in BA order Arabic speakers.

### 3. Discussion

Results indicate that productions were considerably more native-like in the immediate post-test compared to the pre-test. In **Figure 1**, the blue line indicates

**Table 3.** Comparison of AB and BA performance.

Test Order	Pre-Test	Imm. Post-Test	Del. Post-Test
AB Order	49%	90%	84%
BA Order	43%	95%	80%

**Figure 1.** Percentage of native-like productions for each participant per pre-test and immediate post-test.

the pre-test and the red line indicates the immediate post-test. Native-like items in the pre-test were less than half the tokens (46%), but this increased to 92% in the immediate post-test. The pretest had 49 non-native-like productions, but there were only seven in the immediate post-test.

This suggests that the instruction, consisting of a few minutes of practice with the elocution drill (also serving as a mnemonic) together with a small amount of background information, was highly effective for short-term modification of these onset clusters. Given these favorable results, we recommend conducting more extensive research with larger sample sizes to build upon these findings. Nonetheless, the results of this small study at least indicate that brief but carefully constructed elocution drills can be used successfully to modify pronunciation with extremely minimal use by learners.

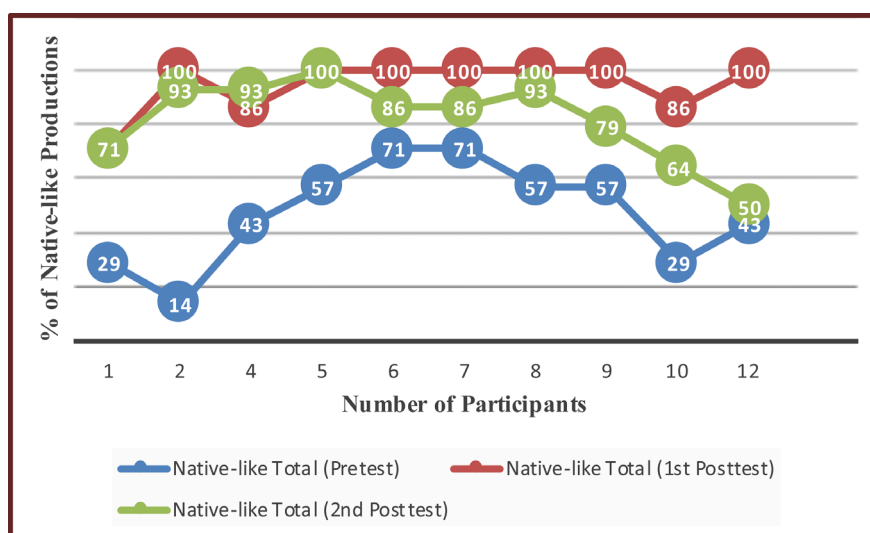
Furthermore, it had been predicted that if participants achieved native-like pronunciation of three-consonant onsets, they might immediately also pronounce two-consonant onsets in native-like fashion. Although the study was not designed to determine whether this occurs with all two-consonant onsets in English, the norm of performance suggests that such an outcome is hopeful. In the pre-test, there were seven non-native-like productions of two-consonant onsets, but in the immediate post-test, there was none. This is a preliminary indication that if learners master marked (three-consonant) onsets, less marked (two-consonant) onsets might also be mastered naturally without additional effort.



**Figure 2** depicts the results of the delayed post-test, compared with the previous tests. Only data from participants who completed all three tests are included. The blue line is the pre-test, the red line is the immediate post-test, and the green line is the delayed post-test. Results of the delayed post-test show that participants still largely maintained the trend toward native-like production. The numbers of native-like productions were 33 of 70 possible (47%) in the pretest, 66 of 70 possible (94%) in the immediate post-test, and 114 of 140 possible (81%) in the delayed post-test. Despite a decrease from the immediate post-test, native-like performance still remained well above pre-test level.

As in the pre-test and immediate post-test, epenthesis, deletion, and feature change occurred in the delayed post-test. Of these, epenthesis was again the most common, occurring 22 times (85%). Deletion occurred three times (11%) and feature change occurred once (4%). All instances of epenthesis occurred after the first consonant of the cluster (i.e., CVCC), unlike in the immediate post-test where the epenthesis was found before the first consonant also. Deletion occurred in two words: “*squad*” and “*stream*”. In the pre-test, “*squad*” exhibited the most deletions (six occurrences). In the delayed post-test, this item was simplified twice by deleting the third consonant [w]. Similarly, the third consonant was deleted by one participant in “*stream*”. Feature change was observed only once by fricativizing the /p/ to [f] in “*spray*”.

Although learner strategies were used less in the two post-tests than in the pre-test, the preferences of learner strategies remained consistent. That is, participants tended to epenthesize rather than delete or change features. Of note, in the delayed post-test, as in the immediate post-test, non-native-like productions occurred with only three-consonant onsets, as there were no non-native-like two-consonant onsets produced. This is again suggestive that the prediction of UG and the Markedness Hypothesis is born out, though further research is necessary to test this comprehensively.



**Figure 2.** Percentage of native-like productions for each participant per test.

## 4. Conclusion

This study was theoretically founded on SAT, UG, and the Markedness Hypothesis. According to SAT, receiving declarative knowledge of English onsets in a format that permits their proceduralization should, with practice, result in their successful native-like acquisition. According to UG and the Markedness Hypothesis, acquisition of marked three-consonant onsets should simultaneously also result in less marked two-consonant onsets being acquired with no additional effort. The results of this small-scale study support, in a preliminary way, these theoretical predictions. This technique provides learners with crucial declarative knowledge packaged in such a way as to be memorable (a mnemonic device) and easy to practice, facilitating proceduralization and automatization. It is also an efficient instructional technique, as it targets only marked onset clusters in the hope that less marked onsets might be acquired without being specifically practiced.

## Acknowledgements

We thank the participants in our study and the faculty at George Mason who assisted us in working with them.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- Anderson, J. R. (1976). *Language, Memory, and Thought*. New York, NY: Lawrence Erlbaum Associates.
- Boersma, P., & Weenink, D. (2016). *Praat: Doing Phonetics by Computer* (Computer Program). <http://www.praat.org/>
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- DeKeyser, R. (1998). Beyond Focus on Form: Cognitive Perspectives on Learning and Practicing Grammar. In C. Doughty, & J. Williams (Eds.), *Focus on Forum in Classroom Second Language Acquisition* (pp. 42-63). New York, NY: Cambridge University Press.
- Foote, J. A., Holtby, A. K., & Derwing, T. M. (2012). Survey of the Teaching of Pronunciation in Adult ESL Programs in Canada, 2010. *TESL Canada Journal*, 29, 1-22. <https://doi.org/10.18806/tesl.v29i1.1086>
- Greenberg, J. (1976). *Language Universals*. The Hague, The Netherlands, NY: Mouton.
- La Cruz, E., & Savaria, L. (2010). Patterns of Regularity in Final Cluster Reduction in L2 English by L1 Spanish Speakers. *Entre Lenguas*, 15, 49-57.
- McLaughlin, B. (1987). *Theories of Second Language Acquisition*. London, UK: Edward Arnold.
- Weinberger, S. (1987). The Influence of Linguistic Context on Syllable Simplification. In G. Ioup, & S. Weinberger (Eds.), *Interlanguage Phonology: The Acquisition of a Second Language Sound System* (pp. 401-417). Cambridge, MA: Newbury House.

## **Appendix A**

### **Form A**

Squad  
Lake  
Split  
Modern  
Scream  
Pick  
Spring  
Flag  
People  
Clue  
Strong

### **Form B**

Spray  
Fly  
Mood  
Square  
Confident  
Splint  
Night  
Screen  
Shelf  
Class  
Stream