Do Adults Spling/Splang?

An Investigation Into the Over-irregularization of Past Tense Conjugations in English

Linguistics 270
University of Pennsylvania

By: Kiawehokua Tarnas, Hannah Gu, and Maria Baez

Abstract

In certain experimental settings, adults may over-irregularize verb conjugations. This paper incorporates a modified experimental set-up of Bybee & Moder's 1983 experiment in order to determine if and how adults deviate from regular conjugation rules for English verbs in the past tense form. We use a Google Form to carry out a rating task among undergraduate participants in which they are presented with a base form of a verb and are asked to decide the naturalness of three past tense forms on a predetermined scale. Further, we review the details of our methodology and discuss the limitations of using a Google Form to gather our data. Additionally, this paper incorporates the work of Schütze and his criticism of the methodology of Wug studies, a predecessor of Bybee & Moder's experiment. We hypothesized that our participants would prefer an "-ed" past tense conjugation for all verb categories (except one) and incorrectly presumed that final consonants would have the greatest effect on the naturalness rating of the past tense forms. Finally, we theorised that there would be a noticeable difference in the numerical ratings that our subjects gave for each option. Our data supported the first point of our hypothesis; however, it did not attribute a causative role to the final consonant in determining the naturalness of a specific conjugation nor did it demonstrate as high a variance in the rating of conjugations as anticipated.

Background

The study of irregularities in morphologically conditioned rules suggests that certain alterations (i.e. vowel changes in English verbs when conjugating the past tense form) are preserved in many languages and, as a result, are patterned and rarely ever random in nature. While children rarely, if ever, overuse an exception in the same way that they may overuse a rule (Maslen, 2004), it is worth noting that specific experimental settings will often create the space for adults to over-irregularize in certain cases. This is exemplified by Bybee & Moder's 1983 study examining the organization of what they identify as the most productive class of strong verbs in the English language: the string/strung English verb class. Bybee & Moder aimed to delineate how verb classes are defined as well as how certain classes will expand once innovations are developed by an individual. Their study builds off of the concept of lexical selection procedure and emphasizes the importance of schemas in defining verb classes. Schemas are defined as lexical associations between existing forms of a category and have been shown to mediate associations among English irregular verbs in the past tense form (Bybee & Slobin, 1982). In defining these categories, Bybee & Moder emphasize Rosch & Mervis 1975's proposal that each category will have a common conjugation with its own characteristic attributes that members of a class will share. These categories are "natural categories," and the morphological classes examined in Bybee & Moder's study as well as our own can be organized according to these very same structures. They suggest that innovations within a verb class are the result of associations made between lexical items themselves according to their schemas, especially if framed in a past tense context such that the morphological function can be accessed when looking to expand verb classes.

This paper uses Bybee & Moder's 1983 experiment as a foundation to investigate how adults may deviate from regular conjugation rules and over-irregularize English verbs in the past tense form. By adapting the methods and stimuli used in Bybee & Moder's study and applying them to a population of undergraduate students, we aimed to identify the past tense conjugations of specific verb categories that were most commonly rated as "natural" by participants. Bybee & Moder found that there is an overwhelming preference for conjugating the past tense form of most verb categories by adding the "-ed" suffix, with the exception of verbs ending in velar-nasal consonants. In our study, we hypothesized that: (a) there exists a preference for "-ed" past tense conjugations for all verb categories excluding those ending with velar-nasal consonants; (b) the final consonants of a verb category play a greater role than the initial consonants in determining the rated difference in naturalness of conjugated past tense forms; and (c) there will be a high variance in the rating of conjugations between our participants.

Subjects

The subjects of the experiment were 54 undergraduate students. All participants were native speakers of English and had no previous experience with the experiment.

Methods

The experiment utilized a list of 32 nonce verbs. These verbs were all monosyllabic and varied phonologically along two dimensions: the initial and final consonant clusters. The initial-consonant segments were of four types: the consonant /s/ followed by a consonant cluster (sCC); the consonant /s/ followed by a single consonant (sC); a two-consonant cluster that did not contain /s/ (CC); and a single consonant (C). Final consonant segments were in the velar-nasal (/ŋk ŋ/); velar (/k g/); nasal (/m n/); or other single consonant (C) categories. All verbs contained the middle /i/ vowel between the two consonant clusters. A list of these verbs appears in the appendix of the paper.

A Google Form was used to carry out the study. Subjects were presented with 32 present tense verbs, each within its own section present on separate pages. Each verb was shown in the present tense and was then followed by the three past tense forms that varied with the conjugations "-ed", /æ/, and /ʌ/. The order of presentation of these three past tense forms was randomized within each section. The subjects were then asked to rate the naturalness of the three past tense forms of the 32 verbs on a scale of 1-7, in which a value of 1 indicated a very unnatural past tense form and a value of 7 indicated a very natural past tense form. The scale was presented as a horizontal row of fill-in bubbles.

The sentence "Sam really likes to … Yesterday he … for 3 hours" was used as a template for the subjects to refer to as they felt necessary throughout the duration of the experiment. The sentence-completion frame and a brief explanation of the scale were repeated in the instructions of each section for the participants' convenience. Subjects were also given an explanation of how this list of 32 verbs functioned. The instructions at the start of the experiment explained that each of these 32 made-up verbs has a different meaning for a newly discovered action. In our

instructions, the nonce word "spling" was used as an example verb for a new and exciting way to bounce a paper clip. The participants were instructed to complete the form in its entirety. In order to control for this, we set up the Google Form such that each question required an answer in order to move on to the following sections. Participants were informed that the experimenters were looking for intuitions and gut feelings, as opposed to looking for what the participant may have learned in school. This experiment did not place a time constraint on the participants.

Results

The data resulting from this experiment was statistically significant and interesting from a linguistic standpoint. This section will consider the data on its own, address its statistical significance, highlight interesting parts of the data, and lay the foundation for comparison to previously produced data that will be done in the Discussion section.

The data is visually represented in a number of bar graphs that can be found in the Figures section. Each graph looked at a specific category of base verb and the mean rating (on a scale from 1 to 7) that each of its past tense conjugations received (Figures 2.1-2.9). A number of the graphs show the expected trend of a noticeably higher rating for the naturalness of the “-ed” conjugation (Figures 2.1, 2.3-2.5, 2.7-2.9). The majority of the graphs also showed very similar average ratings for the /æ/ and /ʌ/ irregular past tense conjugations (Figures 2.1, 2.2, 2.4-2.6, 2.9). However, a few conditions, especially the velar (k/g) final, had notably different ratings between the /æ/ and /ʌ/ irregular past tense conjugations (with /ʌ/ being higher) (Figure 2.7). In contrast, the graph for the velar-nasal final shows extraordinarily close average ratings for all three of the possible past tense conjugations (Figure 2.6). For the “-ed” conjugation, the velar-nasal final condition (Figure 2.6) had the lowest and the double consonant (CC) initial, velar final and nasal final conditions (Figures 2.4, 2.7, 2.8) had the highest average ratings. For the /æ/ conjugation, the velar final condition (Figure 2.7) had the lowest and the velar-nasal final condition (Figure 2.6) had the highest average rating. For the /ʌ/ conjugation, the sCC initial, SC initial, and velar-nasal final condition (Figures 2.2, 2.3, 2.6) had the highest average rating while the other graphs were all lower and around the same mark as each other (Figures 2.1, 2.4, 2.5, 2.7-2.9).

In order to look at the results statistically as well as visually, the t-statistics are shown in the following table (Table 1.1).

Paired T-stats								
	Initials				Finals			
	sCC	sC	CC	C	ŋ/ŋk	k/g	n/m	C
"/æ/" to "-ed"	0.00999669	0.02716376	0.03276263	0.02660896	0.00064955	0.05032784	0.02524027	0.02695466
"/ʌ/" to "-ed"	0.00887567	0.01502721	0.03226873	0.02763369	0.0002598	0.02358106	0.03546707	0.02582396
"/ʌ/" to "/æ/"	0.00154484	0.01211688	0.0004025	0.00062394	0.00044663	0.0248324	0.00954792	0.00148582

Table 1.1 This table contains the paired t-values of initial and final consonants for all comparisons of the three past-tense verb conjugations.

Since each conjugation of each word was rated by every individual participant, the data was able to be paired and paired t-values were made for comparisons between each of the conjugations in every condition. The rows in Table 1.1 are labeled with “X to Y”, indicating that the comparison was done by subtracting the values of Y from X. The t-values represent how significant the closeness of the ratings of each of the conjugations in the comparison are. To point out the statistical representations of some of the visual phenomena that was noted above, the velar (k/g) final had the largest (least significant) t-value representing closeness across all conditions for the comparison between the /æ/ and /ʌ/ irregular past tense conjugations. Also, the lowest (most significant) t-values appeared in the velar-nasal final condition across all conjugation comparisons. Both of these are exactly what would be expected from the visual analysis of the data discussed above. The t-values can be better understood from the frame of the following table (Table 1.2) that shows whether the above t-values are considered significant with a p-value of 0.01.

Significant with p-value of 0.01								
	Initials				Finals			
	sCC	sC	CC	C	ŋ/ŋk	k/g	n/m	C
"/æ/" to "-ed"	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE
"/ʌ/" to "-ed"	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE
"/ʌ/" to "/æ/"	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE

Table 1.2 This table summarizes whether or not the paired t-statistics of initial and final consonants with the past-tense verb conjugations are significant in comparison to a p value of 0.01.

Many of the t-values produced are significant. With a p-value of 0.05, all t-values were considered significant except for one: the comparison from "/æ/" to "-ed" for the velar (k/g) final. The p-value of 0.01 was used for a higher threshold and in order to have more obvious patterns

of statistical significance across conjugation and word categories: The "/ʌ/" to "/æ/" condition consistently had the most significantly close ratings. The sCC initial and the velar-nasal final both consistently had significantly close ratings across all conjugation comparisons. None of the other word categories had as significantly similar enough ratings between their "-ed" condition or either of the other conditions. This is because of the notably higher rating that the "-ed" conjugation received in these conditions. This result is inline with our hypothesis and the data that has previously been collected on this topic. Interpretation of these data and their implications will be addressed at the end of the Discussion section.

Figures

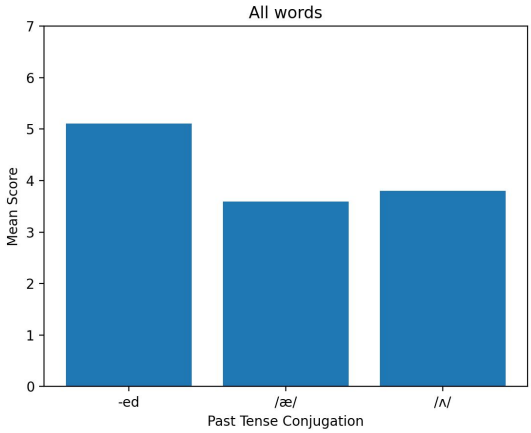

Figure 2.1
This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations and all words.

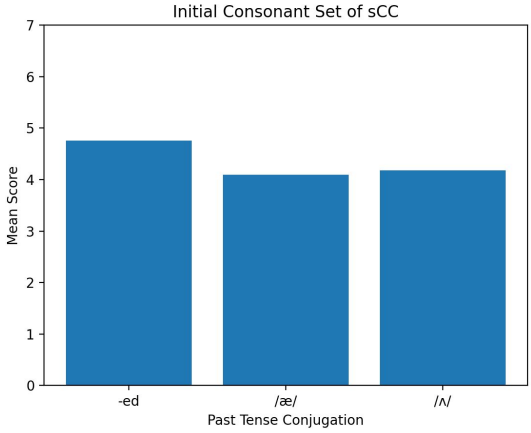

Figure 2.2
This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the initial consonant set of sCC.

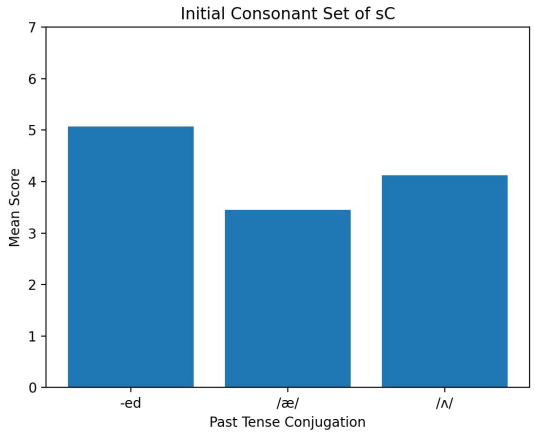

Figure 2.3

This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the initial consonant set of sC.

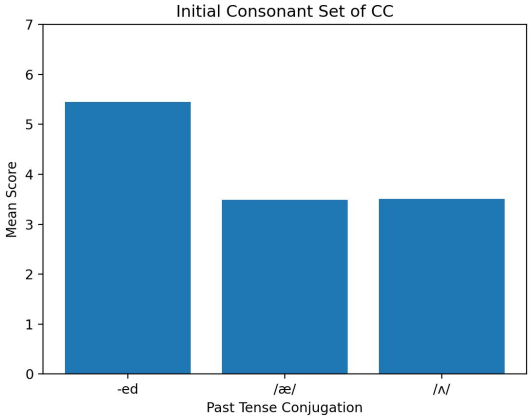

Figure 2.4

This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the initial consonant set of CC.

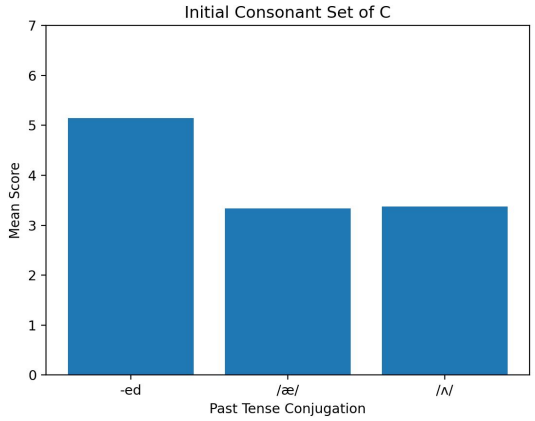

Figure 2.5

This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the initial consonant set of C.

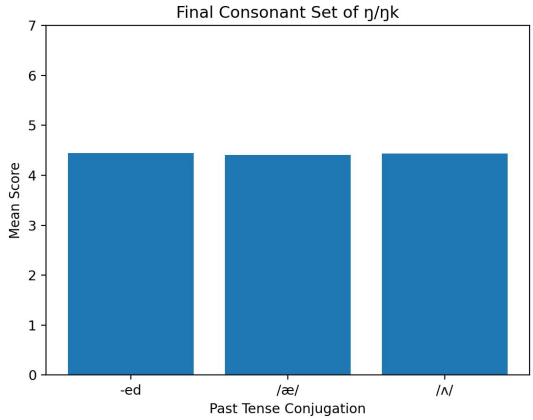

Figure 2.6

This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the final consonant set of ŋ/ŋk (velar-nasal).

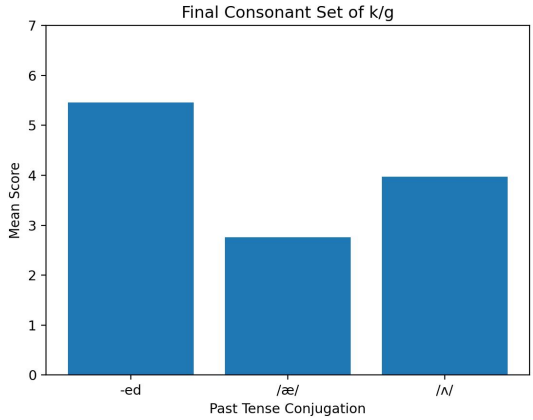

Figure 2.7

This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the final consonant set of k/g (velar). .

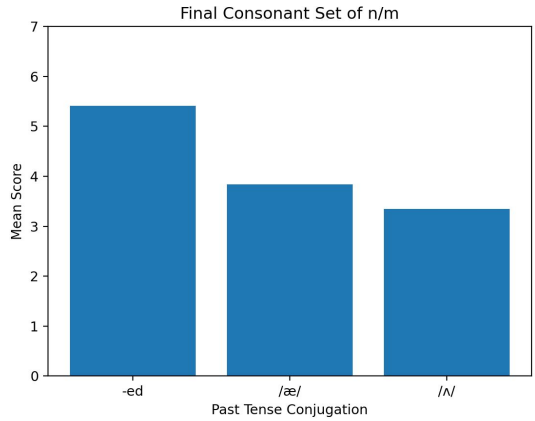

Figure 2.8

This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the final consonant set of n/m (nasal).

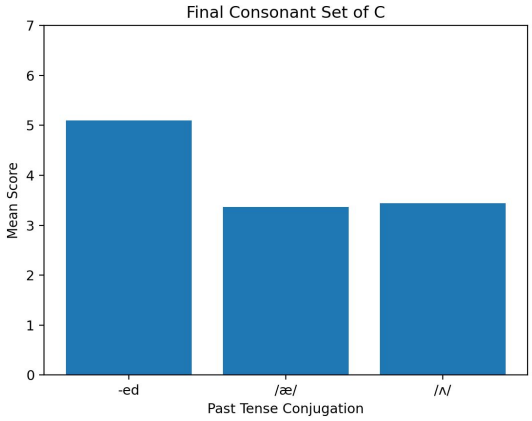

Figure 2.9
This bar graph summarizes the mean score of participants' naturalness rating for all three past-tense conjugations with the base word having the final consonant set of C.

Discussion of Methods

The experiment only tested along two dimensions: the initial consonants and the final consonants. We would have liked to account for the vowel, but given our interest in the initial and final consonants, both of which Bybee & Moder stated had a greater effect on the results than the vowel, we chose to focus on the other parameters and use monosyllabic verbs only containing the /i/ vowel. This is shown in the word list of the Appendix. Given the time restraints of this study and lack of incentive for participants, there were constraints on data-collection. This study focused on the attributes that were deemed more important by previous research. Specifically, Bybee & Moder noted that the vowel makes a significant difference only when other features of the base verb are present. The number of nonce verbs was scaled down from 93 nonce verbs in the model experiment by Bybee & Moder to 32 nonce verbs in our version of the experiment.

A limitation of using Google Forms was that we were unable to prevent participants from returning to previous sections to change their answers. However, we anticipate its effects on our results would be negligible; it is unlikely that participants went back to review and change their responses to this 32 section long Google Form as there was no compensation provided for participating. Therefore, participants had no motivation to go back to edit previous answers. We also chose not to impose a time constraint on our participants unlike Bybee & Slobin, who had their participants supply the past-tense form of the verb that they heard as quickly as possible. Although Bybee & Moder presented the verbs orally, due to the time frame of the experiment and the decreased likelihood that we would obtain willing participants to commit to a timed experiment session over Zoom, we opted for an untimed Google Form to reach more subjects.

The objective of introducing the information about each of these verbs having different meanings for newly discovered actions was to address the neologism scenario as discussed by Schütze. We chose to expand on this issue by identifying the 32 nonce verbs as each being made up for newly discovered actions. He differentiates between the dictionary scenario, in which

these nonce verbs are an obscure verb that could be found in an English dictionary, and the neologism scenario in which a nonce verb is used to describe a new action. Schütze argues that few adult Wug studies, which our reference experiment stems from, use instructions that hint at which scenario is intended by the investigators. The intention of the investigators, according to Schütze, could have a serious impact on how participants chose to approach the task. This scenario can cause the participant to proceed through the neologism assimilation strategy, in which they assume that they are in an everyday neologism situation and must consider how these new words function in their language. An alternative course of action is the neologism acquisition strategy, defined as the strategy in which participants follow the special grammar rules specific to the novel word. This strategy attempts to simulate normal vocabulary acquisition when a word is added to their lexicon, similar to encountering a new word during language acquisition as a child (2005).

The past tense forms, which can be found in the Appendix, were made by using “-ed” and changing the vowel (/æ/ and /ʌ/). We did not give the subjects the option to change the vowel and add “-ed”, as Bybee & Moder's study only found that a few participants came up with this kind of response. The /æ/ and /ʌ/ were the only vowel change forms offered as they were the most common vowel changes among subjects in Bybee & Moder's study. Additionally, the “-ed” form was changed accordingly in order to preserve the present tense form. For example, the past tense “-ed” form of the verb "kib" became "kibbed" (pronounced kibd) in order to maintain the "kib" sound so the participant did not read "kibed" (pronounced kaibd). This was important because we did a nonverbal experiment and needed to account for how subjects might mentally pronounce the past tense “-ed” forms, whereas our reference experiment had each of the verbs presented orally.

No real verbs were included among the nonce verbs, unlike the Bybee & Moder reference experiment in which they utilized 16 real verbs mixed in with their 93 nonce verbs. Schütze pinpointed this as favoring the dictionary scenario. If real words were interspersed throughout, participants would think through their experience with the English language and model their choices off of words that they already know. Schütze explains that people would assess the nonce verb and its similarity to existing verbs using properties that may be relevant to the behavior of the verb, just based on their past experience. They might have done that regardless of whether or not we presented the real words in the experiment itself, but choosing not to include real words was a preventative measure to decrease the odds that they would do so based off of the Google Form's content.

Bybee & Moder used a forced choice task in their study by asking participants to provide a single past-tense form of the verbs that they were presented with. Our study did not use forced choice and instead had participants rate the naturalness of three different past tense forms, choosing to follow Ullman's acceptability rating style (Schütze, 2005). This decision appealed to how "right" subjects felt the past-tense forms were but reduced the pressure of having to choose a form that seemed to be the most "right" or natural. While a certain form may not have been one that the participant themselves would use, having a scale present for each of the possible

conjugations allowed participants to account for forms that may be considered natural or commonly used by others. Furthermore, the rating task was used in order to ensure that there was no coercive influence that caused subjects to favor the traditional "-ed" past tense, a fault of the elicitation version of the Wug test (Schütze, 2005).

Discussion of Results

Despite the differences between the methodology of Bybee & Moder and the methods of our study, the data collected can be compared in a number of ways. Bybee & Moder presented their data in data tables with percentages for each conjugation which represented how often people chose that conjugation as the most natural. Recall that this was a forced choice experiment and participants could only provide one conjugation. The differential between the percentages they got for each conjugation can be compared to the differential between conjugations shown on the bar graphs in the Figures section.

This method of comparison can be utilized to evaluate our hypothesis. Three things in particular were hypothesized: It was predicted that there would exist a preference for "-ed" past-tense conjugations for all verb categories excluding those ending with velar-nasal consonants based on the results of Bybee & Moder's experiment that found exactly this. The results in this paper mirror this effect. Visually, all graphs show obvious preferences for an "-ed" past tense conjugation except for Figure 2.6 on the velar-nasal final. It was also predicted that the final consonants of a verb category would be more impactful on the conjugation ratings than the initial consonants because this effect was reported in Bybee & Moder's paper. Our results did not reflect this prediction. Table 1.2 clearly demonstrates the similarity between the impact of initials and finals on the rated conjugation naturalness of each category. Finally, it was also predicted that there would be high variance in the rating of conjugations between our participants. This prediction had no basis in Bybee & Moder's results because the forced choice task did not allow for this comparison; therefore, the prediction was based on logical consideration of the task rather than any previous data. Surprisingly, there was not as much variance as predicted. Of course, there was still notable variance, in that most words had ratings that ranged almost the entire spectrum (from 1 to 7), but the data was more clustered than expected. This may have been due to the larger-than-expected sample size. We had 54 participants, each of which completed the task for 8 words in each category. This adds up to 432 total trials per category.

There exists a number of similarities between the data from each paper that was not mentioned in the hypothesis. An especially notable similarity between the results is that, for the "-ed" conjugation, the velar-nasal final condition received the least amount of support as a natural conjugation while the CC initial and velar final conditions all received the largest amount of support as a natural conjugation. This finding was quite consistent across the two studies. Both studies were also consistent with the result that the /ʌ/ conjugation was treated as less natural than the /æ/ conjugation in most conditions. Even the exceptions to this rule were consistent between studies: the sC initial, CC initial, and velar-nasal conditions had very similar naturalness scores for the /ʌ/ and /æ/ conjugations. A final similarity between the two studies was

that the velar-nasal condition was the least divergent across all conditions and the velar condition was the most.

There were also a few differences between the two papers' results as well. The most important of these was the difference in how extreme the "-ed" form displayed given preference in Bybee & Moder's paper. The bar graphs in the Figures section show noticeable differences, but nothing as extreme as a 63% to 12% differential as was found in the CC condition between the "-ed" and /ʌ/ conjugations in Bybee & Moder's data. There were also a few minor differences between the results. For instance, while Bybee & Moder found that the C final condition produced almost twice the approval for /æ/ than /ʌ/ conjugations, Figure 2.9 shows the two conjugations as being nearly identical (with the /ʌ/ conjugation even being slightly higher). There was also a similar difference in the C initial condition between the papers. Bybee & Moder found that the /æ/ conjugation was picked as the most natural conjugation 23% of the time while the /ʌ/ conjugation was picked as the most natural conjugation 14% of the time. This large discrepancy was not seen in Figure 2.5, which shows the average rating scores for these two conjugations in the initial C condition. The scores were again nearly identical with the /ʌ/ conjugation even being slightly higher.

Conclusion and Implications

These data provide a number of implications. Firstly, forced choice experiments do indeed bias the data towards the regular conjugation of "-ed"; Bybee & Moder's use of a forced choice task resulted in data that showed a much higher preference for the regular conjugation. The other differences between the data are minor and sparse enough that their implications may not be reliable. Nevertheless, the similarities between the datasets are useful. Despite differences in methodologies, the similarities between the two of them show greater validity in the experimental results of both experiments. This is especially important as the data demonstrates the existence of real effects created by the change of initial and final consonant clusters on the naturalness of a specific past tense conjugation. This is the most important implication of these results. Changes in consonant clusters, specifically categories like the velar-nasal final, had significant impacts on how participants felt about their different conjugations.

The results of our study are noteworthy and give way to future work in the field of linguistics that could examine the external validity of our results. There is a concern that the experimental setting in which these data were collected may have altered the results. An analysis of a large set of speech recordings from outside the environment of an experiment could be done to determine if the experimental setting has significant effects on the data. Further research could also be done on different categories of verbs. Although this experiment, for the sake of precision, focused on 8 categories of verbs, there is room for improvement. Future studies could explore the various other categories of initial and final consonant clusters and focus on manipulation of middle vowels. This study utilized previous data and criticisms of past experiments to produce results that could serve as a stepping stone to our better understanding of this topic.

Appendix

Categories:	initial_consonants	final_consonants	Word	"-ed"	"/æ/"	"/ʌ/"
1	sCC	ŋk/ŋ	skring	skringed	skrang	skrung
	sCC	ŋk/ŋ	strink	strinked	strank	strunk
2	sCC	k/g	strick	stricked	strack	struck
	sCC	k/g	strig	strigged	strag	strug
3	sCC	n/m	skrim	skrimmed	skram	skrum
	sCC	n/m	sprin	sprinned	spran	sprun
4	sCC	C	spriv	sprivved	sprav	spruv
	sCC	C	skrit	skritted	skrat	skrut
5	sC	ŋk/ŋ	sming	sminged	smang	smung
	sC	ŋk/ŋ	spink	spinked	spank	spunk
6	sC	k/g	smick	smicked	smack	smuck
	sC	k/g	skig	skigged	skag	skug
7	sC	n/m	spim	spimmed	spam	spum
	sC	n/m	stin	stinned	stan	stun
8	sC	C	stid	stidded	stad	stud
	sC	C	smip	smipped	smap	smup
9	CC	ŋk/ŋ	tring	tringed	trang	trung
	CC	ŋk/ŋ	krink	krinked	krank	krunk
10	CC	k/g	trig	trigged	trag	trug
	CC	k/g	glick	glicked	glack	gluck
11	CC	n/m	krin	krinned	kran	krun
	CC	n/m	plim	plimmed	plam	plum
12	CC	C	trib	tribbed	trab	trub
	CC	C	clid	clidded	clad	clud
13	C	ŋk/ŋ	shink	shinked	shank	shunk
	C	ŋk/ŋ	ving	vinged	vang	vung
14	C	k/g	gick	gicked	gack	guck
	C	k/g	sig	sigged	sag	sug
15	C	n/m	tim	timmed	tam	tum
	C	n/m	vin	vinned	van	vun

16	C	C	sid	sidded	sad	sud
	C	C	kib	kibbed	kab	kub

References

Bybee, J., & Moder, C. (1983). Morphological Classes as Natural Categories. *Language, 59*(2), 251-270. doi:10.2307/413574

Bybee, J., & Slobin, D. (1982). Rules and Schemas in the Development and Use of the English past Tense. *Language, 58*(2), 265-289. doi:10.2307/414099

Maslen, R. J., Theakston, A. L., Lieven, E. V., & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of speech, language, and hearing research : JSLHR*, *47*(6), 1319–1333. https://doi.org/10.1044/1092-4388(2004/099)

Rosch, E., & Mervis, C. (1975). Family Resemblances: Studies in the Internal Structure of Categories. Cognitive Psychology, 7(4).

Schütze, Carson T. "Thinking About What We Are Asking Speakers to Do." *Linguistic Evidence Studies in Generative Grammar* (2005): 457-84. Print.