

## Measuring the Relationship of Structure to Use: Determinants of the Extent of Recycle in Repetition Repair\*

VSEVOLOD M. KAPATSINSKI  
*University of New Mexico*

### 0. Introduction

When a speaker produces a disfluency, defined as an interruption in the flow of speech, he/she may often repeat a part of the utterance already produced. When making such repetition repairs, speakers may repeat only the last word they (partially) produced, as in (1) and (2), or they may choose to go back (recycle) further, repeating two, three, or even more words. The dependent variable used in this study is the Extent of the Recycle in Repetition Repair (ERRR), which is equal to the number of words the speaker repeats, where the final word he/she has produced is counted as one unit, regardless of whether it was completed. Thus, examples (1) and (2) have ERRR=1, example (3) has ERRR=2, and example (4) has ERRR=3. Equivalently, we can say that the repair goes back to the first transition (word boundary) in (1) and (2), the second transition in (3), and the third transition in (4). We will call the part in bold the repairing, and the part in italics the repaired, and we will say that repair goes back to *appreciated* in (1), *just* in (2), *areas* in (3), and *I* in (4), i.e. the word the repair goes back to is the word preceding the repaired. The location of the plus sign in (1)-(4) is the location of the repair onset. All of the examples are taken from the Switchboard Corpus (Godfrey et al, 1992).

- (1) I really appreciated [*the*, + **the**] whole, uh, English class
- (2) I have been involved, uh, just [*slightly*-, + **slightly**] in a case
- (3) The crime level is not as high as it is in other areas [*of the*, + **of the**] city
- (4) I [*had a similar*, + **had a similar**] health plan

The extent of recycle in repetition repair has been used as evidence for (hierarchical) constituent structure (e.g. DuBois, 1974; Fox and Jaspersen, 1995; Clark and Wasow, 1998; Ford et al, 2003; Sanchez-Ayala, 2003): people tend to

---

\* My special thanks go to Joan Bybee, Jill Morford, Caroline Smith, Rena Torres-Cacoulllos, and Catherine Travis for helpful and insightful comments on earlier versions of this paper.

recycle back to the nearest constituent boundary. However, Fox and Jasperson (1995) located a major exception to this generalization: people tend not to recycle back to the beginning of the VP.

A probabilistic or frequency-based account of constituent structure (e.g. Gregory et al, 1999, Jurafsky, 2002, Bybee 2002, 2003) would account for this finding by noting that the transitional probability for any given verb after any given subject is higher than the transitional probability of any direct object after any given verb or that mutual information (the frequency of the two-word string divided by the product of the frequencies of the component words) or simple string frequency of subject + verb collocations is higher than mutual information or string frequency of verb + object collocations.

This paper surveys the full range of online frequency-based and probabilistic (FP) factors that could predict the extent of recycle in repair and explicitly spells out the implications of different formulations. We apply a corpus-based multivariate method to determine whether an FP factor or syntactic constituency is a better predictor of the extent of the recycle. The results indicate that speakers do tend to go back to the nearest constituent boundary except when that boundary is high in backward transitional probability.

## **1. Methods**

### **1.1. Sampling and Exclusions**

Four samples were taken from a tagged for disfluencies 1,000,000 word subset of the Switchboard Corpus, which consists of two-party telephone conversations between native American English speakers on predetermined topics (Godfrey et al, 1992): 1) a random sample of 250 repairs with ERRR=1 that were chosen using a random number generator available at [www.random.org](http://www.random.org); 2) a random sample of 250 repairs with ERRR=2; 3) a sample of 47 repairs with ERRR=3, which contained all such repairs available in the corpus; and 4) a completely random sample of 950 repairs, which reflects the proportions of one, two, and three-transition repairs in the corpus.

Several types of repairs were systematically excluded from the samples.

First, replacement repairs (Sanchez-Ayala, 2003) were excluded. These are repairs, in which the repairing is not an exact repetition of the repaired but replaces some part of the repaired. The reason for this exclusion is that in these cases the speaker has to repeat the word following the word being replaced for the hearer to know how the repairing fits in with the preceding utterance. Thus, the choice of ERRR is not free.

In addition, repetitions that are separated from the utterance they repeat by other lexical material are excluded because there is no way to know that we are dealing with the same phenomenon in these cases, which may well be accidental or emphatic repetitions rather than repairs, and thus could have distinct properties.

The largest set of excluded repairs consisted of repairs that were initiated within three words from a clause boundary as in (5)-(7). The reason for this exclusion is that the tendency to recycle back to the nearest clause boundary

overrides all of the clause-internal factors we are interested in. When a clause boundary is in the second transition preceding repair onset, speakers recycle to it in 92 out of 100 tokens of such repairs we have collected, compared to the average recycling rate to second transition of 18%<sup>1</sup> and hence including such tokens would result in floor effects for other factors and obscure the differences between them. It is assumed that there is a clause boundary after clause-connecting conjunctions.

- (5) I feel kind of sorry for them right now, because the people [are, + are, + are,] uh, wanting things
- (6) Because [I often, + I often] sort of wonder how, having never been to the Soviet Union, um, how different the people there really are,
- (7) I mean, [this, + this] is amazing

Also excluded were tokens initiated within three words of the completion of another repair as in (8). The reason for this exclusion is to avoid priming effects and speaker-specific preferences.

- (8) I was really familiar [with a lot, + with a lot] [of, + of] the AOR type music.

If two or more repair tokens were initiated after identical three-word sequences, only the first token found was included so that effects obtained cannot be attributed to particular lexical items or collocations sampled.

Another exclusion consists of repairs within three words of a contracted form, such as 've, 's, 'd, n't, of in *kind of/sort of/lots of*, and *to* in *going to*. This exclusion is due to difficulties in calculating the number of transitions due to phonological reduction. One cannot start the repairing with a contracted word but there *is* a (low-frequency) possibility to recycle back to the beginning of the contracted element and start the repairing with its full version. Thus these tokens are ambiguous as to how many transitions they contain. The low likelihood of a recycle to the beginning of the contraction may be due to the high frequency of the collocation or the fact (resulting from the high frequency of the combination) that the contracted element has to be changed for the repairing to be grammatical. Therefore, including such cases would bias the sample towards FP factors. By excluding these tokens, we ensure that our findings are unambiguously attributable to FP factors, rather than phonological reduction.

The range of values for ERRR examined in this study is restricted to one, two, and three-word repairs. This decision is motivated by the fact that all tokens with ERRR greater than four were recycles back to a clause boundary, which we have excluded from the study because of our interest in clause-internal constituent

---

<sup>1</sup> The tendency to recycle back to the clause boundary is apparently not universal and is reported to be absent in Japanese, cf. Ford et al (2003).

structure. There were only four tokens of such repairs in the corpus, three with ERRR=5, and one with ERRR=6. There were three tokens with ERRR=4 that were not recycled back to a clause boundary. All of these tokens include a frequent collocation (*a lot of* in (9), *bring'em over* in (10), and *for the sake of* in (11)), which is often phonologically fused and hence does not include possible transitions before *of* in (9) and (11) and before *them* in (10). These tokens were excluded from consideration because of the presence of transitions at which the repairing cannot be the same as the repaired. However, the fact that all extra-long repairs include a frequent collocation provides support for FP factors.

- (9) I do [a lot of the, + uh, a lot of [the, + the ba-, + the back] work, and stuff.  
 (10) I get a whole bunch of them [and bring'em ov-, + and bring'em over.  
 (11) I'm sure they would all of a sudden band together just [for the sake of, + for the sake [of, + uh, of, + of] unity against us or something

## 1.2. Factors

### 1.2.1. Structural Factors

Two structural factors are considered in this study: 1) NCB, defined as the tendency to recycle back to the nearest constituent boundary, which has been the only factor proposed in the literature as a predictor of ERRR; and 2) OPEN, defined as the tendency to recycle back to the nearest open class item, where open-class items include nouns, verbs, adjectives, and adverbs.

The following set of constituents is used in this study: NP, VP, PP, *to* VP, AP, and Conj. The VP boundary within the *to* VP is excluded from the set of constituent boundaries as is the NP boundary within the PP.

While the *to* VP construction (12) has traditionally been assumed to be a purpose clause, these clauses lack subjects, and their initial boundaries pattern like clause-internal constituent boundaries and not like clause boundaries with respect to repair. Thus *to* VP may be better considered a type of a prepositional construction of the form P VP, similar to the P NP structure of PP.

- (12) They wanted some gray [to do the, [[+]] to do the] trim.

The reason for the exclusion of the P XP internal boundaries from consideration as attractors of the recycle is the definition of *constituent* we adopt. Levelt (1989) has argued that, since, given unidirectional binary branching, every word boundary is a constituent boundary, the statement that repair tends to go back to the nearest constituent boundary is equivalent to saying that it tends to go back to the nearest transition. Thus, Levelt proposed that the repaired and the repairing must be conjoinable with *or*. However, this constraint prohibits any repairs that were initiated after determiners as in (13), a common repair.

- (13) Sometimes I get bogged down [in the, [[+]] {D you know,} in the ] ideology (\*in the or in the<sup>2</sup>)

We adopt a definition that is more semantically-oriented by defining a constituent as a phrase that has or may have a syntactic function where the syntactic functions are subject, object, oblique, and verb.

### **1.2.2. Relative vs. Absolute Frequency and Probability**

FP factors can be divided in terms of whether they involve relative or absolute frequency or probability. Relative FP factors state that the repair goes back to the nearest transition with a lower frequency or probability than the neighboring transitions. Absolute FP factors state that the repair goes back to the nearest transition with a frequency or probability lower than some numerical value. While relative FP factors assume an online comparison of transitions by the speaker, absolute FP factors involve no online computation and refer only to long-term properties of collocations. Both types of FP factors are used in this study. Absolute FP factors will be denoted by absolute value bars around the factor name, e.g. |SF| stands for absolute string frequency.

### **1.2.3. Frequency-Based Factors**

Usage-based grammaticization theory claims that words used together fuse together (Bybee, 2002, 2003). Five frequency-based factors are coded in this study: 1) relative two-word string frequency (SF); 2) absolute two-word string frequency (|SF|); 3) relative mutual information (MI), 4) absolute two-word mutual information (MI); and 5) recursive string frequency (RSF).

String frequency is the number of times a certain two-word combination appears in the corpus. Mutual information, on the other hand, is string frequency divided by the frequencies of the words in the string as shown in eq. (1), (2), where F is frequency, w is a word, and n is its number in the production sequence.

eq. (1):  $SF = F(w_n; w_{n\pm 1})$

eq. (2):  $MI = \frac{F(w_n; w_{n\pm 1})}{F(w) * F(w_{n\pm 1})}$

While |SF| postulates that speakers will tend to recycle back to a transition with a two-word frequency lower than some predetermined numerical value, RSF, which is also an absolute FP factor, uses strings of different length for different transitions. Thus, the recycle goes back one more transition if the frequency of the string that begins with the word immediately preceding the transition and ending with the word immediately preceding repair onset is larger than a prespecified numerical value (a.k.a. the cutoff point). Thus, while in first transition before the

---

<sup>2</sup> There are no instances of *the or* on Switchboard, except for one repair (or the + or the)

repair onset the RSF string consists of two words, it consists of three words in second transition as shown in (14), where *of the* is the RSF string in first transition, while *areas of the* is the RSF string in second transition. The cutoff point is shared by all of the strings of a particular length, e.g. the median SF of two-word strings is larger than the median SF of three-word strings.

(14) The crime level is not as high as it is in other [areas [of the]] + of the city

The frequencies of the words and strings were taken from the corpus. All phonologically identical uses of a word were grouped together as long as all of these uses belonged to the same lexical category. Thus the frequency of *like* in *John and Mary like swimming* would not include the tokens where *like* is a conjunction or a discourse marker. Similarly, the demonstrative *that* was separated from the conjunction *that*. On the other hand, all nominal uses of *bank* would be grouped together, regardless of whether they meant ‘a financial institution’ or ‘shore’. All phonologically distinct forms of a word were considered separately, except for phonologically conditioned alternations. Thus, the frequency of *likes* was calculated independently of the frequency of *like* but *an* and *a* were grouped together.

#### 1.2.4. Probabilistic Factors

Jurafsky and his collaborators (Gregory et al, 1999; Jurafsky, 2002) have proposed that frequency is just a special case of probability and that items that are more predictable on the basis of the item that preceded them in the processing sequence tend to reduce and fuse with the preceding item. Thus, in this model, the emergence of constituent structure is inherently directional and hearer-based, i.e. it depends on the speaker’s estimate of the predictability of the next item for the hearer.

In the case of repair, the processing sequence need not correspond to the temporal order in which words are produced. Rather, speakers may attempt to model the hearer’s comprehension process. Repetition repairs involve a syntactic reanalysis on the part of the hearer, in which, according to serial models of reanalysis (e.g. Fodor and Inoue, 1998), hearers move in the direction opposite to the direction of speech production and word-by-word through the sentence. Therefore, we may consider it likely that the speaker’s processing sequence in determining ERRR is the opposite of the temporal sequence of producing the utterance. Thus, backward transitional probability (TP-B) may a priori be considered as plausible a determinant of ERRR as direct transitional probability (TP-D). Transitional probability is defined as the number of times a given two-word string occurs in the corpus divided by the number of times the first (TP-D) or the second (TP-B) word occurs (eq. 3, 4). Both absolute and relative TP-D and TP-B are used in this study.

$$\text{eq. (3): } TP - D = \frac{F(w_{n-1}; w)}{F(w_{n-1})}$$

$$\text{eq. (4): } TP - B \pm \frac{F(w; w_{n-1})}{F(w_{n-1})}$$

Example (15) presents a token from the corpus and the application of the relative TP-D and TP-B to it. W4 is the word in or immediately after which the repair is initiated, w3 is the immediately preceding word, and so on. The SF of the second and third transitions are equal in this example. However, TP-D ((9/other) ? (9/areas))<sup>3</sup> predicts ERRR=3, since *areas* is less frequent than *from*, hence TP-D is smallest in the third transition (w1-w2), while TP-B ((9/areas) ? (9/of)) correctly predicts ERRR=2, since *of* is more frequent than *areas*, hence TP-B is smallest in second transition (w2-w3). The arrow shows actual ERRR.

(15) The crime level is not as high as it is in other areas [of the, + of the] city.

String	String Frequency	TP-D	TP-B	Transition
other areas w1 w2	9 F <sub>w1</sub> =1033;F <sub>w2</sub> =98	0.0087 = 9/1033	0.0918 = 9/98	3
areas of w2 w3	9 F <sub>w2</sub> =98;F <sub>w3</sub> =12457	0.0233 = 9/98	0.0007 = 9/12457	2 ←
of the w3 w4	1890 F <sub>w3</sub> =12457;F <sub>w4</sub> =2000 7	0.1517 = 1890/12457	0.0945 = 1890/20007	1

### 1.2.5. Magnitude-Dependent Dominance

Magnitude-dependent dominance is a way to combine factors, where by combining factors we mean considering their joint effect.

For magnitude-dependent dominance to be used, at least one of the component factors must be able to make predictions with various degrees of certainty. This is not true of the two structural factors in our study, since a constituent boundary either is or is not there. However, it is true of the FP factors: a relative FP factor predicts that speakers will recycle to the transition that has the lowest frequency or transitional probability. The magnitude of the difference between transitions varies from token to token. Similarly, absolute FP factors stipulate that the recycle will go to the nearest transition with frequency or probability lower than some value, but the magnitude of the difference between the frequency or probability in a transition and the cutoff value varies from transition to transition. Therefore, magnitude-dependent combinations of FP factors with structural factors are possible and are considered in this study. Under magnitude-dependent dominance, the prediction of the composite factor is the same as the prediction of the factor with higher magnitude in the token under consideration. If one of the factors has

<sup>3</sup> where '?' stands for '>, <, or ='

constant magnitude, e.g. NCB, a cutoff value is specified for the variable magnitude factor it is combined with, e.g. TP-B. Then TP-B is overridden by NCB whenever TP-B's magnitude is smaller than the cutoff. Magnitude (delta) of an FP factor is calculated through eq. (5),(6).

$$\text{eq. (5)} \Delta F_i = \text{MAX}((M/X),(M/Y),(M/Z))$$

where F is a factor; i is a token number; X, Y, and Z are factor values in the three transitions preceding repair onset; and

$$\text{eq. (6)} M = X+Y+Z-\text{MIN}(X,Y,Z)-\text{MAX}(X,Y,Z)$$

For magnitudes of different factors to be compared, they must be reduced to a common denominator since factors vary in their average magnitudes, in the range of the magnitudes, and in how magnitudes are distributed across that range. We ranked each factor's magnitudes separately from other factors from the largest (#1) to the smallest (#547)<sup>4</sup>. Then the factor with the larger magnitude is the factor whose magnitude has the smaller ranking number, i.e. is closer to the maximum magnitude *for that factor*. Magnitude-dependent dominance can achieve significant improvement in performance over the best component factor even when the component factors are highly correlated and even when they make predictions on all tokens, as long as the factors are less confident in their wrong predictions than they are in their correct ones.<sup>5</sup>

### 1.2.6. Cutoff Adjustment

Cutoff adjustment can be used on FP factors. Cutoff adjusted relative FP factors do not stipulate that the recycle will go back to the transition with the lowest frequency or probability but specify a relation between the frequency or probability in that transition and its nearest competitor as in eq. (7), where k is a constant coefficient specified by the factor, x is a transition, and X, Y, and Z are factor values in the three transitions considered.

$$\text{eq. (7): choose } x \text{ iff } X < k \cdot \text{MIN}(Y,Z)$$

---

<sup>4</sup> The number of magnitudes ranked corresponds to the number of repair tokens in the sample.

<sup>5</sup> Two other ways to combine factors are possible: 1) strict dominance and 2) equality. Under strict dominance, when the component factors make competing predictions, the prediction of the factor designated as dominant is the prediction of the composite factor. The less dominant factors have their say only when the more dominant factors make no prediction. This is not optimal for our task because only OPEN makes no predictions in more than 10% of the tokens. Therefore, a strictly dominant combination of factors will not perform better than the dominant factor. Under equality, when the component factors make competing predictions, the composite factor makes no prediction. The utility of equality in two-factor models is limited, hence equality is not used here.

We must note that any  $k$  different from 1 is an assumption about which transition the speaker defaults to. If  $k$  for a transition is smaller than one, the speaker prefers that particular transition to others, and the reverse holds for  $k > 1$ . Since this means that changing  $k$  from 1 means introducing preferences for transitions as an independent factor, which could only be justified if recycles to some transition were less predictable than recycles to others, we always use  $k=1$  for relative FP factors. However, cutoff adjustment is used with absolute FP factors where no implication of preferences for transitions holds.

### **1.2.7. Window of Attention**

Finally, we consider two possibilities for the window of attention the speaker uses while recycling. One possibility (serial recycle) is that the speaker considers transitions in pairs beginning with the transition that is closest to repair onset and only goes back if needed. The other possibility (instant recycle) is that for every token the speaker simultaneously considers all three transitions preceding repair onset as possible sites to recycle to. In other words, the question is whether the speaker compares the first transition to the third transition, or chooses the first transition if its frequency or probability is lower than that of the second transition without regard to the third transition. A related issue is whether in serial recycle for absolute FP factors, the speaker may return to a transition after rejecting it based on comparison with cutoff when the frequencies and probabilities of the following transitions are also larger than cutoff. We assume that this does not occur, and the speaker chooses the last transition he/she considers, but this is an area for future work. Absolute FP factors can only produce many predictions with serial recycle and thus are not considered under instant recycle.

## **2. Results**

In the random sample of 950 repairs, great majority have  $ERRR=1$  (Table 1). Factors are not compared over this sample because of the unequal distribution of tokens across  $ERRR$  values. If factors were compared in this sample, factors performing well at  $ERRR=1$  would win over factors performing well at  $ERRR=2$  and  $ERRR=3$ . Since we are interested in predicting the distribution, this would not be appropriate.

Table 1: Distribution of tokens across  $ERRR$  in random sample of 950 repairs

$ERRR$	1	2	3
% of data	79%	18%	3%

Factor performance is defined as the percent of tokens on which a factor makes the correct prediction, out of the total number of tokens on which it makes a prediction. Note that 33% is the chance level. The results are shown in table 2. Shaded cells contain factors that are not significantly different from the best factor at the .05 level according to the chi-square test and hence are not eliminated. The cutoff for all of the absolute FP factor combinations has been optimized, while the

magnitude-dependent combinations shown are ones that give a significant improvement over the best component factor. The ideal factor (combination) would not be eliminated in any of the data sets and would successfully describe the distribution function, i.e. would be equally predictive regardless of ERRR. The need for this additional condition is that it ensures that the winning factor is independent and need not be combined with a preference for certain transitions.

Table 2: Factor performance in balanced sample

a. Serial Recycle

ERRR	TP-B	TP-D	SF	MI	NCB	OPEN	TP-B	TP-D	SF	MI	RSF	NCB/ TP-B
1	63%	38%	48%	62%	74%	62%	70%	61%	83%	43%	50%	72%
2	68%	27%	53%	49%	69%	55%	61%	28%	37%	59%	79%	70%
3	34%	19%	19%	30%	44%	32%	60%	28%	43%	62%	37%	57%

b. Instant Recycle

ERRR	TP-B	TP-D	SF	MI	NCB	TP-B/NCB	OPEN
1	47%	15%	37%	42%	74%	63%	62%
2	68%	27%	53%	49%	69%	76%	55%
3	70%	32%	57%	62%	44%	64%	32%

From these data, we note that NCB is the only factor that is not eliminated under either serial or instant recycle. However, it performs relatively poorly in predicting long (ERRR=3) recycles (44%), which motivates combining it with |TP-B| for serial recycle and TP-B for instant recycle. We may note that the strengths of TP-B and NCB lie in different samples in that NCB is good at predicting ERRR=1 tokens and bad at predicting ERRR=3 tokens. The reverse is true of TP-B. Their magnitude-dependent combinations provide the correct distribution function. Since magnitude-dependent combinations perform well only when the FP factor on average has a larger magnitude when it makes a correct prediction than when it makes an incorrect one, good performance of NCB/TP-B indicates that speakers recycle to NCB unless the transition it favors is exceptionally high in TP-B (relative to some other transition).

The various factors examined are correlated to different degrees. FP factors that perform well tend to be correlated with NCB more than those that perform poorly (table 3).

Table 3: Factor correlations and mean percentages of data covered across samples

Factors	Absolute FP Factors		Factors	Relative FP Factors	
	Correlation	Coverage		Correlation	Coverage
NCB/ TP-B	.66	64%	NCB/TP-B	.73	58%
NCB/ SF	.55	54%	NCB/MI	.64	49%
NCB/ MI	.45	55%	NCB/SF	.58	44%
NCB/ TP-D	.38	39%	MSB/TP-D	.46	26%

### *Determinants of Recycle Extent in Repetition Repair*

Thus, the best frequency or probability-based factors, the absolute and relative TP-B are highly correlated with NCB. How might we explain this? The reason for this correlation is that English is a preposing language, therefore in the sequence shown in (16), the constituent boundary is at transition 1 where the lexical item precedes the grammatical item.

(16) lex] 1 [gram 2 lex]

Since the frequency of a lexical item is on average lower than the frequency of a grammatical item, TP-B of transition 1 (where the frequency of the gram is in the denominator) would be lower than TP-B of transition 2 (where the frequency of the lexical item is in the denominator). The reverse is true of TP-D. Therefore in preposing languages TP-B serves as a reliable predictor of syntactic constituent boundaries, while TP-D may serve this function in postposing languages.

Extending this work to postposing languages is likely to provide us with a valuable glimpse into the relationship between syntactic constituency and various FP factors. If the good performance of TP-B reflects the processing sequence involved in repair, it should perform well regardless of whether a language is preposing or postposing. If, on the other hand, its good performance is due to its correlation with NCB, we expect it to perform poorly in postposing languages and TP-D to perform well.

Another direction for future research is to include prosodic structure as a possible factor. In addition, it is not yet clear whether all constituent boundaries are equally attractive for repair to go back to, and what predicts how attractive a given boundary (type) is. More research on the window of attention in repair as well as reanalysis is needed. Finally, the method used here may, with slight modifications, be applied to other dependent variables, such as disfluency onset locations, phonological fusion of frequent collocations, and codeswitching sites.

I hope to have demonstrated a rigorous quantitative method to examine the relationship between structure and use in syntax, which can be applied to other variables that are supposed to be sensitive to structural and FP factors. The fact that a magnitude-dependent combination of NCB and TP-B works indicates that speakers do tend to recycle to the nearest constituent boundary rather than a transition with lower frequency or probability by default whereas long recycles are motivated by an exceptionally high TP-B (relative or absolute) at transitions closer to repair onset. This is tied to the observation that recycles with  $ERRR > 3$  that do not go back to a clause boundary consist solely of recycles that include very frequent collocations. Speakers seem to ignore differences in frequencies and probabilities between words in a sentence, unless they are larger than a particular value. This value is lower than what is necessary for phonological reduction and fusion, since all tokens of phonological fusion were excluded from the sample. It appears that backward transitional probability influences the constituent structure of an utterance even when no effect on phonology can be observed. Other FP

factors do not appear to be reliable predictors of constituent boundaries or the extent of recycle in repair.

## References

- Bybee, J. 2002. Sequentiality as the basis of constituent structure. In T. Givon and B. F. Malle (eds.) *The Evolution of Language out of Pre-Language*, 109-134. Amsterdam, Philadelphia: John Benjamins.
- Bybee, J. 2003. Cognitive processes in grammaticization. In M. Tomasello (ed.) *The New Psychology of Language, Vol. 2*, 145-67. Mahwah, London: Erlbaum.
- Clark, H. H. and T. Wasow. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37: 201-42.
- DuBois, J. W. 1974. Syntax in mid-sentence. *Berkeley Studies in Syntax and Semantics*, 1: III.1-III.25.
- Fodor, J. D. and A. Inoue. 1998. Attach anyway. In J. D. Fodor and F. Ferreira (eds.) *Reanalysis in Human Sentence Processing*, 101-142. Dordrecht: Kluwer.
- Ford, C. E., B. A. Fox, and S. A. Thompson. 2003. Social Interaction and Grammar. In M. Tomasello (ed.) *The New Psychology of Language, Vol. 2*, 119-44. Mahwah, London: Erlbaum.
- Fox, B. A. and R. Jasperson. 1995. A syntactic exploration of repair in English conversation. In P. W. Davis (ed.) *Alternative Linguistics: Descriptive and Theoretical Modes*, 77-134. Amsterdam/Philadelphia: John Benjamins.
- Godfrey, J. J., E. C. Holliman, and J. McDaniel. 1992. SWITCHBOARD: Telephone Speech Corpus for Research and Development. *IEEE ICASSP*: 1517-1520.
- Gregory, M. L., W. D. Raymond, A. Bell, E. Fosler-Lussier, and D. Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *CLS* 35.
- Jurafsky, D. 2002. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, and S. Jannedy (eds.) *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA, London: MIT Press.
- Sanchez-Ayala, I. 2003. Patterns of repair and constituency in Spanish conversation. Paper presented at the Hispanic Linguistics Symposium 6, Albuquerque, NM.

Vsevolod Kapatsinski,  
MSC03-2130, Linguistics,  
1 University of New Mexico,  
Albuquerque, NM, 87131.

alator@unm.edu