Korean Journal of English Language and Linguistics
Vol. 19, No. 4, Winter 2019, 629-643
10.15738/kjell.19.4.201912.629

Frequency and Acoustic Reduction in English *-ment* Derivatives

Jae-Hyun Sung
(Yonsei University)

Sung, Jae-Hyun. 2019. Frequency and acoustic reduction in English *-ment* derivatives. *Korean Journal of English Language and Linguistics* 19-4, 629-643. This study investigates the influence of frequency on the production of bimorphemic words, and considers which frequency measure is most apt to explain the differences. Previous studies have reported that frequent words are produced faster and more casually than infrequent ones, and that medial segments will yield shorter durations. The present study examines the relation between frequency and the duration of medial segments in English derived words by conducting a production experiment with 6 native speakers of American English using 74 English *-ment* derivatives, and evaluates whole-word frequency, base frequency, and relative frequency (whole-word frequency divided by base frequency) against one another as predictors. The results show that models incorporating any of the three frequency measures strongly predict medial segment duration. Among the three frequency measures, whole-word frequency explained the most variance, across all consonant types. The duration of segments in highly frequent words tends to be shorter than that in relatively infrequent words. Overall, this study confirms that speakers are sensitive to the extralinguistic information associated with the words such as frequency, and in this case, traditional frequency measures (whole-word and base frequencies) are better predictors than relative frequency.

Keywords: English *-ment* derivatives, whole-word frequency, base frequency, relative frequency, acoustic reduction

1. Introduction

It has been widely acknowledged that highly frequent words are produced faster and more casually. Numerous studies on a variety of languages have reported that high lexical frequency generally predicts acoustic reduction and greater coarticulation. Acoustic reduction in high-frequency words can be seen in shorter segmental duration (Pluymaekers et al. 2005 (Dutch), Gahl 2008 (English), Myers and Li 2009 (Taiwan Southern Min)), and greater coarticulation has been demonstrated by greater degree of palatalization (Bush 2001 (English)), voice assimilation (Ernestus et al. 2006 (Dutch)),

and tonal merger (Myers and Li 2009).

Pluymaekers et al. (2005) examined the acoustic duration of four inflectional affixes in Dutch that attach to stems with various frequencies, and higher frequencies of the carrier words resulted in significantly shorter duration of the affixes, represented by shorter duration of the whole affixes, individual segments, or both. In the similar vein, Myer and Li (2009) pointed out that the degree of syllable contraction in Taiwanese disyllabic words is strongly correlated to the lexical frequency, in which higher lexical frequency leads to the greater degree of contraction. Moreover, Bush (2001), Ernestus et al. (2006) and Myer and Li (2009) show that coarticulation and assimilation processes are also heavily influenced by lexical frequency. Segmental sequences in highly frequent words yield a greater degree of coarticulation and assimilation. For instance, speakers tend to palatalize the /j/-preceding consonant more frequently when it is from high-frequency words, e.g., /d/ in *would you* (high frequency) vs. *good year* (low frequency) (Bush 2001), and listeners tend to perceive complete voice assimilation across adjacent segments found in high-frequency words (Ernestus et al. 2006). The findings from these studies, along with a myriad of studies to date, demonstrate that lexical frequency plays a critical role in speech production and perception, in a way that segments at various phonological and morphological boundaries are produced and perceived differently according to their frequency.

One of the recent claims in the literature pertaining word frequency is that highly frequent bimorphemic words tend to be processed like monomorphemic words, known as "whole-word access" (Hay 2002, 2003). For instance, /n/ from the word *government* tends to undergo complete deletion among native speakers of English, whereas /n/ from the word *discernment* rarely does. Hay (2002, 2003) and Kim (2009) claim that relative frequency, as opposed to lexical frequency or absolute frequency (henceforth whole-word frequency), can predict the presence or absence of /n/ deletion in the aforementioned two words. Table 1 shows the frequencies of two English *-ment* derivatives, *government* and *discernment*, retrieved from the Corpus of Contemporary American English (henceforth COCA, Davies 2008), and Figure 1 presents the acoustic representations of those two words obtained in this study. As shown in Table 1, *government* shows very high relative frequency in which its whole-word frequency (i.e., frequency of *government*) exceeds its base frequency (i.e., frequency of *govern*). In contrast, *discernment* exemplifies low relative frequency in which its whole-word frequency (i.e., frequency of *discernment*) is markedly smaller than its base frequency (i.e., frequency of *discern*). Such difference in relative

frequency between two words seems to be linked to massive acoustic reduction represented by complete deletion of /n/, as illustrated in Figure 1.

Table 1. Frequencies of English Words *government* and *discernment*

Words	Frequency	Bases	Frequency	Relative Frequency
government	7693	govern	230	33.448
discernment	4	discern	45	0.089

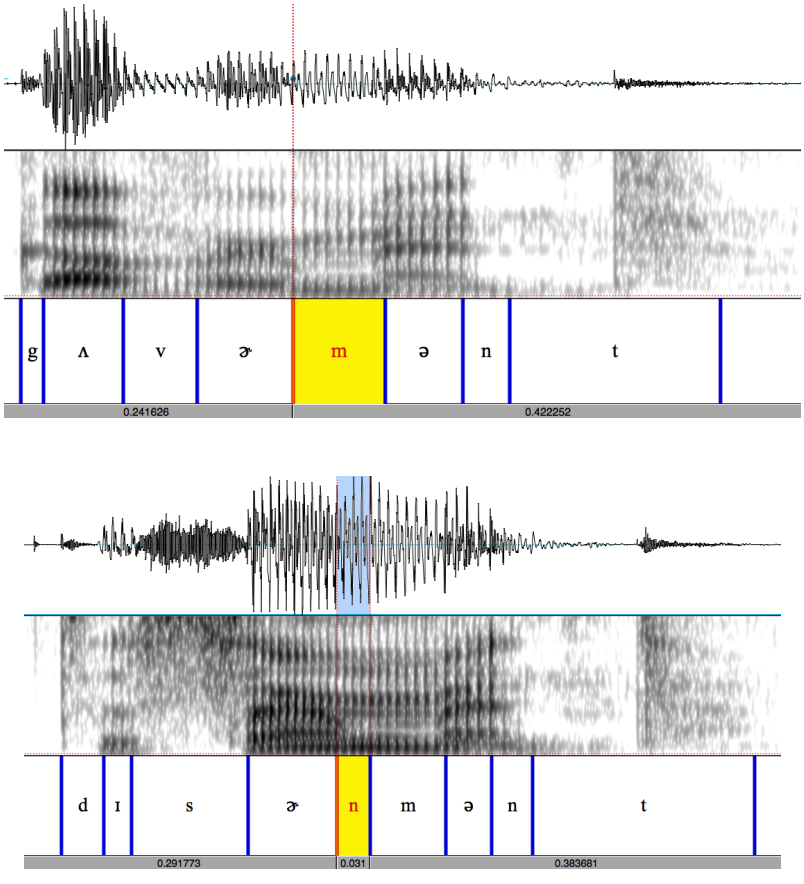

Figure 1. Casual Production of *government* (top) and *discernment* (bottom)

What Table 1 and Figure 1 tell us about word frequency is that various frequency measures might be at play in speech production. As noted in Hay (2002, 2003) and Kim (2009), the two words *government* and *discernment* differ markedly in relative

frequency, but also in whole-word and base frequencies, in that *government* shows higher relative, whole-word, and base frequencies than *discernment* does. This poses an important question on which frequency measure serves as the strongest predictor in boundary segments. For this reason, the present study examines frequency effects by three frequency measures: whole-word frequency, base frequency, and relative frequency. The acoustic characteristics of boundary segments, which are final segments of base words in this study, would reveal whether the same frequency effect would be predicted by these three frequency measures.

This study also investigates whether native speakers of English use the whole-word strategy for highly frequent bimorphemic words in production, and if so, how the whole-word strategy manifests in final segments of base words (e.g., /n/ in *government*). If highly frequent bimorphemic words are perceived as a whole, the acoustic cues at morphological boundaries, to be represented by duration of final segments of base words in this study, may not be as important as those of infrequent bimorphemic words. For this reason, the acoustic cues are expected to be weaker than those of infrequent words, resulting in shorter duration of boundary segments.

2. Methods

2.1 Participants

Six native speakers of American English in their early 20s who lived in Tucson, Arizona at the time of data collection were recruited for the experiment. All the participants were undergraduate students at the University of Arizona.

2.2 Stimuli

This study chose English *-ment* derivatives to examine the acoustic reduction in words with various frequencies. The derivational suffix *-ment* is one of the well-known level 2 suffixes which do not cause stress shift to the base that it attaches to (Katamba and Stonham 2006). As *-ment* has no phonological effect on the base it attaches to, known as one of the 'neutral' affixes, English *-ment* derivatives are suitable in examining the boundary segments in morphologically complex words.

Seventy four English bimorphemic *-ment* words derived from verb bases were chosen

from COCA for stimuli, as well as 54 fillers not containing *-ment*. Previous studies on English cited in this paper used CELEX (Baayen et al. 2001) which represents British English. As this study analyzes the production data from speakers of American English, frequency data was retrieved from COCA which represents American English.

Among many English *-ment* derivatives found in COCA, the present study focuses on English *-ment* derivatives whose verb bases end in an obstruent phoneme (/p, t, d, v, s, z, ʃ, ʧ, ʤ/). Boundary segments were limited to obstruents in this study due to the difficulty in measuring the acoustic duration of sonorants that immediately precede another sonorant.

Table 2. Examples of Test Words and their Frequency Measures

Target Cs	Words	Freq	Bases	Freq	Relative Freq
/p/	development	115,318	develop	95,981	1.201
/t/	statement	49,354	state	54,922	0.899
/d/	commandment	2,018	command	7,287	0.277
/v/	improvement	22,019	improve	47,080	0.959
/s/	reinforcement	3,031	reinforce	10,513	0.288
/z/	advertisement	4,083	advertise	5,139	0.795
/ʃ/	nourishment	767	nourish	1,578	0.486
/ʧ/	attachment	2,937	attach	16,733	0.295
/ʤ/	judgment	23,138	judge	25,656	0.902

Examples of the stimuli are shown in Table 2. Base verbs which the *-ment* suffix attaches to vary in stress and number of syllables[1]. All frequency counts are the sum of all the possible conjugated forms of words. For instance, the frequency of *development* is the sum of instances of *development* and *developments*. The frequency

[1] Stress and number of syllables were not controlled, which might have contributed to the durational patterns in this study. The present study did a post-hoc analysis on the role of stress in the *-ment* derivatives and the role of stress on the duration of boundary segments was not statistically significant, at least for 74 *-ment* derivatives used in this study.

Base words in *-ment* derivatives used in this study vary, and also are not skewed in terms of number of syllables, as shown in the appendix. The role of number of syllables is not of a particular interest in this study, and as base words with various lengths were equally used in different frequency groups, the present study assumes that the number of syllables did not contribute to the durational patterns.

of *develop* is the sum of frequencies of *develop*, *develops*, and *developed* (including both past and past-participle), and *developing* (as gerund).

2.3 Procedure

Participants were instructed to read each word in two carrier sentences "Over the phone she said _____. I said _____ again." on a computer screen. This resulted in two instances of each word, both of which were measured and analyzed. The sentences were recorded in a soundproof booth through a high-quality head-mounted microphone.

2.4 Measurements

Durations of the final obstruents of base verbs, henceforth referred to as boundary segments, were measured. The measurement was based on the offset of F2 of preceding vowels, the offset of voicing from preceding vowels, and abrupt change in acoustic energy shown in waveforms and spectrograms. Figure 2 presents a sample measurement of the boundary segment /t/ in *investment*.

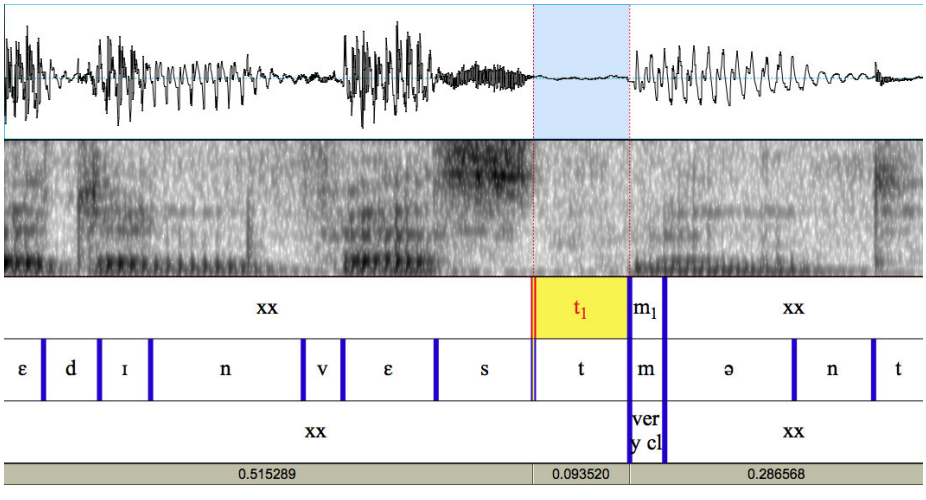

Figure 2. Sample Measurement of the Boundary Segment /t/ (highlighted) in *investment*

2.4 Statistical Analysis

Linear regression was used to evaluate the effects of three frequency measures on the duration of boundary segments, and the frequency measures were log-transformed for statistical analysis. The observations that were considered as outliers, whose log frequencies are smaller than 1 or greater than 6, were excluded from the analysis.

3. Results

3.1 Whole-word Frequency

The regression utilizing whole-word frequency and consonant type significantly predicts segment duration (R^2 = .5764, p < .001) in which higher frequency tends to coincide with shorter segment duration. Moreover, a significant interaction between whole-word frequency and segment type was found. As shown in Figure 3, the negative correlation between segment duration and whole-word frequency, represented by a right diagonal regression line, was found in stops (r^2 = .0415, p < .001) and fricatives (r^2 = .0697, p < .001), but not in affricates (r^2 = .0007, p > .05). The results from stops and fricatives support the hypothesis of this study that highly frequent bimorphemic words show weaker acoustic cues in boundary segments, resulting in shorter segment duration in this study.

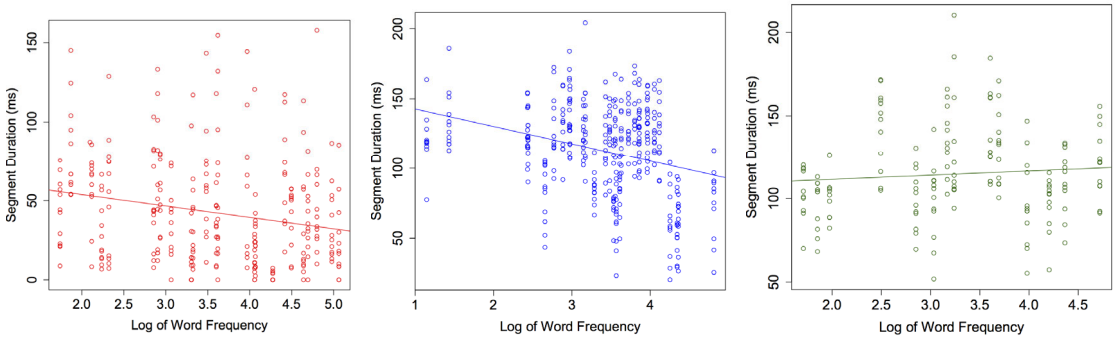

Figure 3. Duration of Boundary Segments (ms) and Whole-word Frequency Measures: Stops (red), Fricatives (blue) and Affricates (green)

3.2 Base Frequency

The regression in which base word frequency was employed as the frequency measure also predicts medial consonant duration (R^2 = .5602, p < .001). That is, higher frequency of base verbs tend to co-occur with shorter duration of boundary segments. Again, significant interaction between base frequency and segment types was found. As illustrated in Figure 4, the negative correlation between segment duration and base frequency was only found in stops (r^2 = .0206, p < .01) and fricatives (r^2 = .0465, p < .001), but not in affricates (r^2 = −.0042, p > .05). The results suggest that base frequency is a similarly apt predictor of medial segment duration.

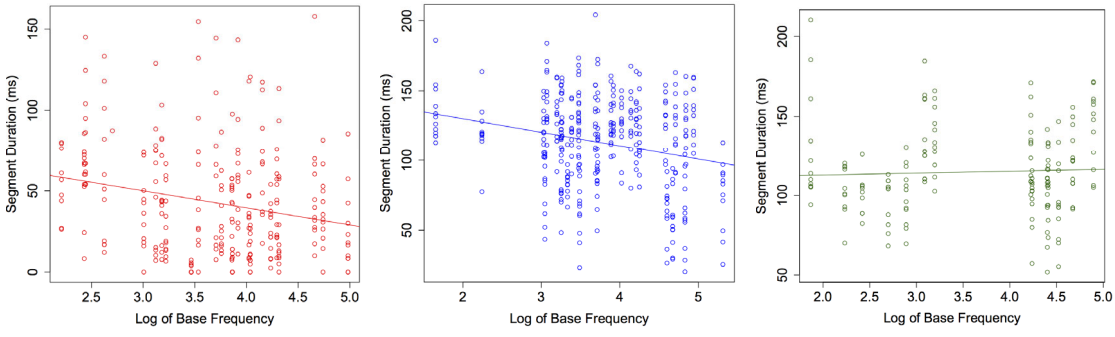

Figure 4. Duration of Boundary Segments (ms) and Base Frequency Measures: Stops (red), Fricatives (blue) and Affricates (green)

3.3 Relative Frequency

A significant model fit was also achieved predicting segment duration with relative frequency and consonant type, shown in Figure 5 (R^2 = .5499, p < .001). By this model, only fricatives show significant correlation with relative frequency (r^2 = .0080, p < .05), but not stops (r^2 = .003, p > .05) nor affricates (r^2 = −.0047, p > .05).

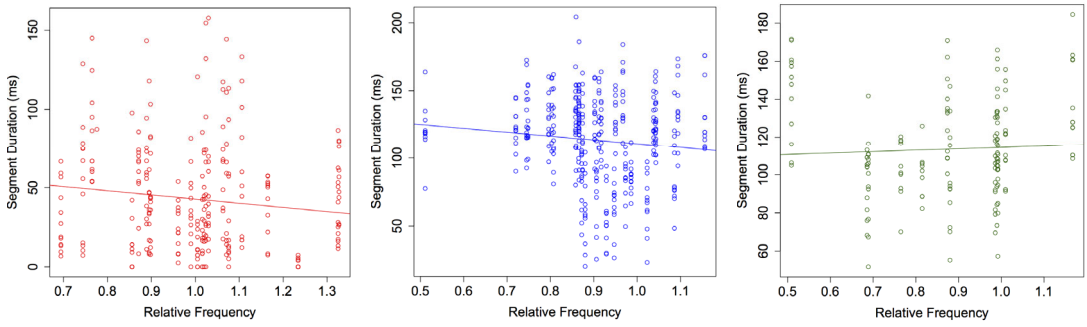

Figure 5. Duration of Boundary Segments (ms) and Relative Frequency Measures: Stops (red), Fricatives (blue) and Affricates (green)

3.4 Summary

Table 4 summarizes the aforementioned statistical results with R^2 scores. As illustrated in the table, whole-word frequency appears to be the strongest predictor of all three consonant types, and base frequency comes as the next strongest one. Relative frequency is the least decisive factor. Contrary to the findings in Hay (2002, 2003) and Kim (2009), relative frequency played a significant role only in duration of target fricatives. Furthermore, none of the frequency measures is a significant predictor of duration of target affricates.

Table 4. Summary of results

		Consonant Types		
		Stops	Fricatives	Affricates
Frequency Measures	Whole-word	.415***	.0697***	.0007
	Base	.0206**	.04649***	-.0042
	Relative	.0030	.0080*	-.0047

***p < .001, **p < .01, *p < .05

4. Discussion and Conclusions

Despite a relatively small number of subjects and test items, the results from the production experiment indicate that frequency effects in English derivatives vary across

different frequency measures. Frequency effects in the *-ment* derivatives were strongest in the whole-word frequency measure, represented by the highest R^2 score, and weakest in relative frequency measure, represented by the lowest R^2 score. The same trend was observed across all consonant types.

The study suggests that high relative frequency is not as strong a predictor of durational reduction in speech production as the previous studies claim (Hay 2002, 2003, Kim 2009). The results from the perception experiment in Hay (2003) and the text-based analysis in Kim (2009) showed that relative frequency of multimorphemic words played a stronger role than whole-word frequency, which was not observed in the current production study. The durational evidence presented in this study implies that frequency effects vary across different modalities of speech or word processing, and the whole word access hypothesis in production might be supported only through whole-word or base frequency, but not relative frequency. This calls for future studies that examine the role of different frequency measures on various speech modalities.

Furthermore, the durational patterns observed in this study exhibit gradience in frequency effects, which has been supported in numerous recent experimental studies that examine frequency-induced phonological processes (Ernestus 2011, Hammond 2004, Hay and Baayen 2005 among many). Along with the general trend that high-frequency lexical items lead to shorter duration than low-frequency ones, the correlation between acoustic duration and different frequencies is quite complex.

Overall, this study investigates the effects of various frequency measures in English derivatives based on an acoustic analysis of *-ment* derivatives. The acoustic data analyzed provide durational evidence for traditional whole-word and base frequencies rather than relative frequencies, and presents a potential mismatch in frequency effects among various aspects of speech/word processing. Future experimental studies are necessary to further examine frequency effects in speech production and perception across various multimorphemic words.

References

Kim, J.-H. 2009. A reconsideration of asymmetries of bracketing paradoxes in English derivation: A corpus-based approach. *English Language and Literature* 55-3, 475-495.

Boersma, P. and D. Weenink. 2011. *Praat: Doing Phonetics by Computer,* Version 5.1.25.

Bush, N. 2001. Frequency effects and word-boundary palatalization in English. In J. L. Bybee and P. Hopper, eds., *Frequency and the Emergence of Linguistic Structure*, 255-280. John Benjamins.

Davies, M. 2008. *Corpus of Contemporary American English. http://www.english-corpora.org/coca/*

Ernestus, M., M. Lahey, F. Verhees and H. Baayen. 2006. Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America* 120, 1040-1051.

Ernestus, M. 2011. Gradience and categoricality in phonological theory. In M. van Oostendorp, C. J. Ewen, E. Hume and K. Rice, eds., *The Blackwell Companion to Phonology: Volume IV. Phonological Interfaces*, 1-22. Wiley.

Gahl, S. 2008. *Time* and *thyme* are not homophones: The effect of lemma frequency on word duration in spontaneous speech. *Language* 84, 474-496.

Hammond, M. 2004. Gradience, phonotactics, and the lexicon in English phonology. *International Journal of English Studies* 4, 1-24.

Hay, J. 2002. From speech perception to morphology: Affix ordering revisited. *Language* 78, 527-555.

Hay, J. 2003. *Causes and Consequences of Word Structure.* Routledge.

Hay, J. and R. H. Baayen. 2005. Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences* 9, 342-348.

Katamba, F. and J. Stonham. 2006. Morphology, 2nd ed. Palgrave Macmillan.

Myers, J. and Y. Li. 2009. Lexical frequency effects in Taiwan Southern Min syllable contraction. *Journal of Phonetics* 37, 212-230.

Pluymaekers, M., Ernestus, M., and R. H. Baayen. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118, 2561-2569.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

Jae-Hyun Sung, Professor
Yonsei University
Department of English Language and Literature
50 Yonsei-ro, Seodaemun-gu
Seoul, Korea
Tel: 02) 2123-2329
E-mail: jsung@yonsei.ac.kr

Received: November 10, 2019
Revised: December 16, 2019
Accepted: December 20, 2019

Appendix

Test Words and Their Frequency Measures

Target Cs	Words	Freq.	Bases	Freq.	Relative Freq.
/s/	abasement	27	abase	45	0.600
/t/	abatement	718	abate	1,513	0.475
/ʤ/	abridgement	51	abridge	171	0.298
/t/	abutment	125	abut	505	0.248
/ʃ/	accomplishment	7,244	accomplish	16,206	0.447
/v/	achievement	22,814	achieve	39,873	0.572
/ʤ/	acknowledgement	1,079	acknowledge	25,183	0.043
/t/	adjustment	11,676	adjust	17,124	0.682
/s/	advancement	3,567	advance	13,826	0.258
/z/	advertisement	4,083	advertise	5,139	0.795
/t/	allotment	857	allot	162	5.290
/z/	amazement	1,932	amaze	2,152	0.898
/d/	amendment	18,803	amend	2,917	6.446
/z/	amusement	3,402	amuse	1,792	1.898
/s/	announcement	9,155	announce	38,200	0.240
/z/	appeasement	453	appease	1,109	0.408
/t/	appointment	11,424	appoint	10,857	1.052
/ʤ/	arrangement	16,015	arrange	17,254	0.928
/s/	assessment	32,297	assess	23,121	1.397
/t/	assortment	2,108	assort	8	263.500
/ʃ/	astonishment	1,398	astonish	1,072	1.304
/ʧ/	attachment	4,937	attach	16,733	0.295
/ʃ/	banishment	275	banish	2,433	0.113
/s/	basement	9,244	base	87,132	0.106
/d/	bombardment	1,160	bombard	1,013	1.145
/ʧ/	catchment	311	catch	78,400	0.004
/d/	commandment	2,018	command	7,287	0.277

/s/	commencement	944	commence	1,829	0.516
/t/	commitment	26,135	commit	14,263	1.832
/t/	comportment	131	comport	269	0.487
/t/	contentment	800	content	421	1.900
/t/	department	95,122	depart	5,718	16.636
/t/	deportment	172	deport	1,661	0.104
/ʤ/	derangement	94	derange	263	0.357
/ʧ/	detachment	1,476	detach	1,572	0.939
/p/	development	115,318	develop	95,981	1.201
/p/	elopement	74	elope	277	0.267
/s/	embarrassment	4,260	embarrass	3,007	1.417
/t/	enactment	1,158	enact	6,236	0.186
/s/	endorsement	5,270	endorse	8,254	0.638
/ʤ/	engagement	9,552	engage	33,177	0.288
/s/	enhancement	2,686	enhance	17,773	0.151
/s/	enticement	271	entice	1,791	0.151
/p/	equipment	31,342	equip	7,231	4.334
/p/	escapement	55	escape	19,689	0.003
/ʃ/	establishment	13,061	establish	51,127	0.255
/t/	excitement	9,308	excite	5,066	1.837
/s/	harassment	6,322	harass	2,982	2.120
/ʧ/	impeachment	4,033	impeach	1,226	3.290
/ʤ/	impingement	71	impinge	499	0.142
/v/	improvement	22,019	improve	47,080	0.468
/t/	incitement	210	incite	1,323	0.159
/t/	indictment	4,590	indict	3,503	1.310
/s/	inducement	589	induce	5,228	0.113
/ʤ/	infringement	709	infringe	776	0.914
/t/	investment	43,780	invest	20,629	2.122
/v/	involvement	17,849	involve	67,548	0.264
/ʤ/	judgment	23,138	judge	25,656	0.902

/ʤ/	management	52,314	manage	47,098	1.111
/v/	movement	64,837	move	205,682	0.315
/ʃ/	nourishment	767	nourish	1,518	0.486
/ʧ/	parchment	1,733	parch	74	23.419
/v/	pavement	3,682	pave	3,072	1.199
/s/	placement	7,173	place	69,420	0.103
/s/	pronouncement	1,478	pronounce	4,883	0.303
/ʃ/	punishment	11,307	punish	7,731	1.463
/ʃ/	ravishment	14	ravish	175	0.080
/t/	recruitment	4,001	recruit	10,558	0.379
/ʃ/	refreshment	932	refresh	1,176	0.793
/s/	reinforcement	3,031	reinforce	10,513	0.288
/t/	resentment	4,169	resent	3,423	1.218
/p/	shipment	3,033	ship	8,260	0.367
/t/	statement	49,354	state	54,922	0.899
/t/	treatment	63,077	treat	45,746	1,379