Using Information Content to Predict Phone Deletion

Keywords: phonology, information theory, phone deletion

Recent years have seen a growing interest in using information theoretic terms to explain linguistic phenomena (Aylett and Turk, 2004; Levy and Jaeger, 2006), as well as a growing availability of large scale phonetic transcription of spontaneous speech (Godfrey and Holliman, 1997; Pitt et al., 2007). This work uses information theoretic terms to better explain the deletion facts found in the Switchboard Corpus, since predicting asymmetries and typologies of deletion in spontaneous speech is essential to our understanding of phonology. The statistics available in Table 1, extracted from the Switchboard Corpus, show an asymmetry in the deletion probability of medial stops in English:

- /ŋ/ is more likely to delete than /n/, /n/ is more likely to delete than /m/
- /t/ is more likely to delete than /k/, /k/ is more likely to delete than /p/
- /d/ is more likely to delete than /b/, /b/ is more likely to delete than /g/.

Current work in phonology cannot explain these asymmetries without arbitrarily stipulating that the saliency of stops in different places varies with regard to voicing and nasality (Côté, 2004), or providing markedness hierarchies for each type of stop (de Lacy, 2002). There have been successful attempts to predict phonological facts using distributional data. Zipf (1929) correctly predicted that more frequent phones are more likely to change, but this claim is not strong enough to predict that /ŋ/ is more likely to delete than /n/, since it is a great deal less frequent in English. Aylett and Turk (2004) and Boomershine et al. (in press) show that the probability of seeing a phone given the previous phone predicts phone deletion and duration, but this is not sufficient to capture large scale phonological properties, e.g. unpredictable /t/s are more likely to delete than predictable /p/s.

I define the basic information value of a phone in context as its contribution to word recognition: the negative log probability of seeing a phone given all the previous phones in the word, $-\log_2(\Pr(\text{phone}|\text{previous phones}))$. This means that the more likely we are to see a phone in a context, the less informative it is in that context. The information value of a phone out of context is taken to be the mean information value of all its occurrences in spoken language. My claim is that this property is *part of the representation of phones*, and influences the typology of phonological processes within a language. When we apply this measurement to English, /ŋ/ and /t/ turn out to be the least informative phones, and we therefore expect to find phonological processes that target them more than other phones. This explains the asymmetries in the deletion of stops in English. In other languages other phones might be less informative and therefore targeted by more phonological processes. /s/ is the least informative phone in Spanish, and we do find /s/-targeting processes in Spanish.

This claim was tested on English, using the Switchboard corpus. I used a logistic regression that controlled for the phonological properties of both the segment and its environment, supra-segmental properties such as its position in a syllable and stress, rate of speech, and previous information theoretic variables used in (Aylett and Turk, 2004) and (Boomershine et al. in press). As Table 2 shows, the mean information value of phones was found to be significant and highly predictive of phone deletion: the less informative a phone is, the more likely it is to delete. After adding this factor, many phonological factors, such as the place of articulation of the deleted phone, did not have any explanatory power.

This experiment provides evidence for the linguistic reality of the information content of phones, their effect on spontaneous speech, and means of discussing why certain phonological phenomena occur while others do not. The information value of phones acts as a driving force in phonology, and yields a principled explanation for phonological typology.

References

Aylett, M. and Turk, A. (2004). Explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47.

Boomershine, A., Hall, K. C., Hume, E., and Johnson, K. The impact of allophony vs. contrast on speech perception. In Avery, E. Dresher, K. R., editor, *Phonological Contrast*. Mouton. In Press.

Côté, M.-H. (2004). Syntagmatic distinctness in consonant deletion. *Phonology*, 21.

de Lacy, P. (2002). *The formal expression of markedness*. PhD thesis, University of Massachusetts Amherst.

Godfrey, J. J. and Holliman, E. (1997). Switchboard-1 release 2. Linguistic Data Consortium, Philadelphia.

Levy, R. and Jaeger, T. F. (2006). Speakers optimize information density through syntactic reduction. In *roceedings of the Twentieth Annual Conference on Neural Information Processing Systems*.

Pitt, M., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., and Fosler-Lussier, E. (2007). Buckeye corpus of conversational speech (2nd release). Department of Psychology, Ohio State University.

Zipf, G. K. (1929). Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology*, 40:1–95.

Table 1: Medial Phone Deletion

Phone	Mean Information Value	Deletion Probability in Switchboard
m	3.0471759	0.02568218
n	1.9025826	0.11142604
ŋ	0.2285215	0.16072842
p	3.9335600	0.02259136
t	1.5667523	0.39027264
k	2.7488146	0.05400620
b	3.9728667	0.07860616
d	1.8870894	0.44070081
g	4.7997063	0.03683492

Table 2: Medial Phone Deletion Model Improvement

Change to model	Residual Deviance	
None: control model	8085.797	
Adding the phone's mean information value	8038.915	-46.88
Adding the next phone's mean information value	8031.067	-7.84
Adding the phone's information value (in its context)	8024.661	-6.40