A New Historical Analysis of /ʌ/

Greg Hanneman
Language Enthusiast and Errant Scribbler
Carnegie Mellon University

Abstract

We introduce a novel analysis of the phoneme /ʌ/, tracing its use from its first appearance in ancient speech to its distribution in modern utterances. We analyze its contemporary distribution with a probabilistic model and provide a number of illustrative examples. We also practice writing in the plural even though there is only one of us.

1 Introduction

Experts generally agree that the first language in the world was the one developed by the prehistoric people of what is now southern France [6], expressed among the famous cave paintings discovered at Lascaux in 1940. The more commonly studied depictions involve hunting and gathering scenes, village feasts, and wars against neighboring tribes. Also among them, however, way back on that one reddish rock just to the left of the *bête noire*, is a fragment of writing that puzzled linguists for decades (Figure 1). It was later analyzed and discovered to be the Diamond Jubilee version of Gregg shorthand [2].

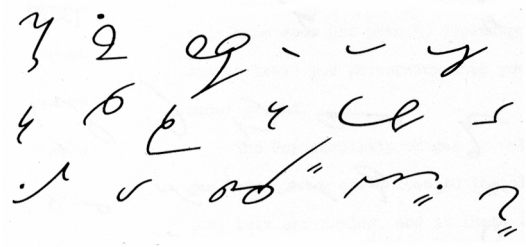

Figure 1: Mysterious fragment of prehistoric writing discovered in a cave at Lascaux.

Of the many phonemes represented in the Lascaux text, one of the most prevalent and useful is /ʌ/. Indeed, the /ʌ/ phoneme's importance, dating from its appearance in well-known prehistoric words like /ʌg/ and continuing to the present day, has been unquestioned for thousands of years. In this paper, we present a novel analysis of /ʌ/, taking especial care to place the use of the sound in its historical and modern contexts. We also describe a series of elicitation experiments designed to provide data on the current use of /ʌ/ in informal speech.

2 Previous Work

Hanneman [3] provided a first-rate introduction of the topic in his very readable recent paper. We cannot hope to improve upon his fine analysis, choosing instead to refer the interested reader to his original work for full details.

3 Experimental Results

The value of /ʌ/ as a space-filler in hesitant and confused speech is well known. Our experiments show that the duration of the sound generally ranges from 550 ms to 1200 ms, with values reaching more than 3000 ms in some speakers. Other so-called "gap" phonemes, such as /a/ or /œ/, are generally less fun to study, so we have no data on them.

3.1 Modern Speech

We do, however, enjoy drawing syntax trees. To handle the appearence of /ʌ/ in a speech stream, we adopt the Penn Treebank annotation UH [4] in each of our examples below. Data taken from two native English speakers between the ages of 23 and 24[1] indicate a somewhat systematic use of UH in casual

[1]The relatively small sample size and homogeneity of our test subjects is due to the fact that they were the only grad students in the lab the night before the SIGBOVIK paper deadline.

speech. A spurious /ʌ/ is most frequenty inserted before a major syntactic constituent (Figure 3.1),

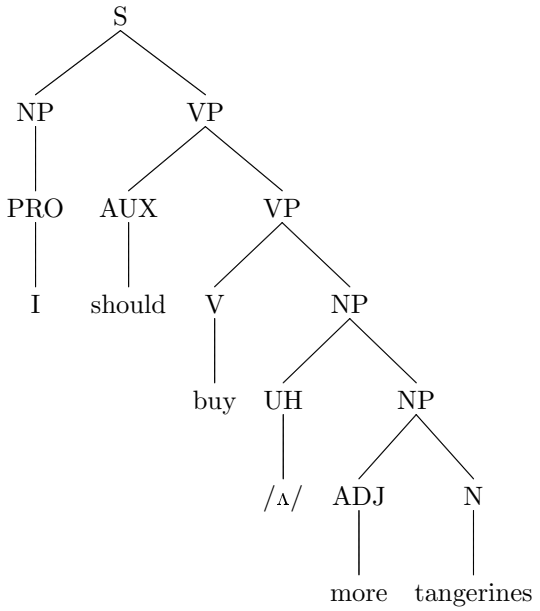

Figure 2: In this example from our corpus, a spurious /ʌ/ occurs before the noun phrase *more tangerines*.

We represent this pre-constituent distribution with a probability model. We explicity model a context-sensitive series of features over lexicalized constituents. In this case, we let c be the syntactic category of the constituent (NP, VP, etc.), h be the lexicalized head word of the constituent (*tangerines*, etc.), and p the current phase of the moon expressed as a real number on the range $[-1, 1]$. We also keep track of a base probability, u, of a /ʌ/ insertion for each speaker s. This gives us the following model:

$$P(u, c, h, p, s) = P(c, h, p)P(u \mid s) \tag{1}$$

The title of this paper, however, promised a *historical* analysis of /ʌ/. Therefore we conclude our discussion of current uses and move on to the next subsection.

3.2 Historical Speech

The earliest wax phonograph cylinders now extant provide a fascinating look at the speech of prehistoric man. (See, for example, Sage [5].) From the beginning, the prevalence of /ʌ/ in some utterances, such as /ʌgə.wʌgə.gʌ/, can approach the level of the letter *e* in modern English texts. These examples show the pervasive extent of the phoneme in ancient speech, and also indicate how a modern player might win at IPA hangman against a Cro-Magnon opponent.

In more modern times, /ʌ/ has continued to be used upon important occasions of all kinds, ranging from Julius Caesar's famous dying words "*Et tu, Brute?* /ʌ/*!*" in 44 B.C. to Queen Victoria's oft-quoted "/ʌ/ , we are not amused." Modern writers, though, who have creative differences with the International Phonetic Alphabet generally prefer to render /ʌ/ as "uh" in normal text; modern copy editors, who insist on regularity of punctuation, set it off with commas. In this form, the sound has shown up in various contemporary contexts, such as "Cake or, uh, death?" and "This is, uh, Spartaaaaa!" The authors feel that further explication beyond these perceptive examples will not be necessary.

4 Error Analysis

Propagation of error or uncertainty can be calculated in two distinct ways [1]. The first, known as the *derivative method*, computes the error in function f due to uncertainty in variable x as:

$$\delta_{fx} = \left|\frac{\partial f}{\partial x}\right| \delta_x \tag{2}$$

Once the error due to each variable in a function is calculated in this way, they are combined via the quadrature method. For a function $f(a, b, ...n)$:

$$\delta_f = \sqrt{\delta_{fa}^2 + \delta_{fb}^2 + \cdots + \delta_{fn}^2} \tag{3}$$

Error can also be calculated using the so-called *computational method*, which is often preferable to first-year physics students because it requires no calculus. Given a function $f(a, b, ..., n)$, the error due to each variable is simply:

$$\delta_{fx} = |f(a, b, ..., x + \delta_x, ..., n) - f(a, b, ..., x, ..., n)| \tag{4}$$

Then the quadrature method of Equation 3 can be applied.

The application of the above formulas to the data is left as an exercise to the reader [10 points]. Write that down in your copybook now.

References

[1] Department of Physics, Case Western Reserve University. *Lab Manual*, 2006.

[2] John Robert Gregg, Louis Leslie, and Charles Zoubek. *Gregg Shorthand*. McGraw-Hill, 1971.

[3] Greg Hanneman. A new historical analysis of /ʌ/. In *Proceedings of SIGBOVIK: Workshop about Symposium on Robotic Dance Party of*, Pittsburgh, PA, April 2007. Association for Computational Heresy.

[4] Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. Building a large annotated corpus of english: The penn treebank. *Computational Linguistics*, 19, 1993.

[5] Glenn Sage. Cylinder of the month. Online, December 1999. http://www.tinfoil.com/cm-9912.htm#e10000.

[6] Ross Steele, Susan St. Onge, and Ronald St. Onge. *La civilisation française en évolution*. Thomson Heinle, 1996.