Chapter 3

Articulatory Effort Reduction

This chapter reports the results of an experiment designed to explore whether articulatory effort reduction plays a role in the lenition of intervocalic stops. Previous studies of articulatory effort, I argue below, are lacking in that they use indirect means to assess effort: either abstract models of the articulatory apparatus, or post-hoc reasoning from the characteristics of the relevant gestures to the relative effort involved. Experiment 1 takes an approach that is, to my knowledge, nearly unique in the literature: it attempts to observe effort reduction in action in the laboratory, by creating conditions meant to encourage subjects to use less effort in speaking.

The premise of the experiment is that a subject who is tired or impaired will produce less effortful articulations than he otherwise would; therefore, by comparing the speech of subjects in normal and tired conditions, we can identify as ‘easy’ those productions that are favored in the tired condition. Of course, saying that we would like to study tired speech is one thing, and actually inducing tiredness in a laboratory setting is another. In Experiment 1, I simulated tiredness
with intoxication, comparing the speech of subjects who were intoxicated with the speech of the same subjects when they were sober. I do not claim that inducing intoxication is the only or even the best way to simulate tiredness in the laboratory; indeed, it is possible that each of the conceivable methods of inducing tiredness or other general impairment (intoxication, exercise, sleep deprivation, emotional stress, etc.) has its own unique characteristics and gives us only a partial picture of speech under impairment. Experiment 1 is intended as an initial exploration of what impairment of some kind might tell us about articulatory effort.

3.1 Assessing Articulatory Effort

Most previous assessments of articulatory effort in the literature fall into one of two categories. In the first type, a researcher argues that some characteristic of a certain type of production causes that production to require more effort than productions that lack that characteristic. The following representative quotations illustrate some of the characteristics of speech sounds that have been claimed to affect their difficulty:

Displacement of articulator “Extreme displacements and extreme velocities are avoided....we find that speech production appears to operate as if physiological processes were governed by a power constraint limiting energy expenditure per unit time.” (Lindblom 1983, 231)

Speed of articulator movement “Matthies et al. (2001) report that the peak velocities of lip movements are greater for clear speech than for conversational speech for /iCu/ syllables. These studies demonstrate that the pro-
duction of clear speech requires more effort and expends more energy than does the production of conversational speech.” (Uchanski 2005, 226)

**Precision of gesture** “Fricative closures require more control and thus more effort.” (Lavoie 2001, 165)

**Stability** “…the coupled-oscillator view on slips of the tongue sees errors as arising from a move towards optimization; they are instances of optimal stability. In this sense, if we start producing a /t/ and a /k/ at the same time instead of alternating between them, we are reducing articulatory effort.” (Pouplier 2003, 2246)

**Tenseness of gesture** “It seems reasonable to posit that [r] is more difficult articulatorily than [l]. This is based on...the inherent tenseness of the trill articulation” (Padgett 2009, 440)

**Energy expended** “…assimilation, defined as reduced distance between two sequentially timed articulatory targets, implies less work per unit time. In a mechanical system such a restructuring of a frequently used sequence of targets will obviously, in the long run, lower energy costs....it does not seem unreasonable to hypothesize by way of analogy that [languages undergo assimilation] to optimize motor control by minimizing physiological energy expenditure.” (Lindblom 1983, 237-238)

Arguments of this type are a good first step in our attempt to understand articulatory effort, but they are far from conclusive. First, it is not always clear whether the observed differences among gestures are really large enough to make a difference. For example, Kingston (2008, 1) argues that lenited segments do not
require less effort than unlenited segments, because the differences between the
two productions in distance traveled by the articulators and in time are extremely
small. Second, the metrics of effort listed above may conflict. For example, frica-
tives require a smaller gesture than stops (suggesting that they require less effort),
but they also require more precision (suggesting that they require more). There
is not always an obvious way to combine these metrics and determine the overall
amount of ‘effort’ required for a given production.

The second way researchers have assessed articulatory effort involves building
an abstract model of the speech apparatus and comparing the production re-
quirements predicted by the model for different segment types. Such approaches
generally interpret one parameter (or more) of the model as a measure of articu-
atory effort, with a given production’s value for that parameter translating into
the degree of effort it requires. Notable examples of models of this type include
the following:

- Lindblom and Sundberg (1971), Lindblom (1983), and Lindblom (1990)
describe a model of tongue shapes and the possible shapes associated with
various consonants and vowels. Lindblom (1983) argues that coarticulation
between consonants and vowels is predicted by a constraint that penalizes
extreme parameter values in the model.

- Westbury and Keating (1986) describe a model of airflow in the vocal tract.
They define ‘easy’ articulations to be those that require the slowest artic-
ulator movements, and show that the model predicts voicing to be easier
in some contexts (intervocally) than in others (initially or finally). The
same model has been applied by Hayes (2004[1999]) to other environments.
• Kirchner (2001b) constructs a mass-spring model of the vocal tract (along the lines sketched in Lindblom (1983, 227-229)). He defines effort as the total force exerted throughout a production, showing that the model predicts that certain types of lenited consonants are less effortful than their unlenited counterparts (for example, that singleton stops are less effortful than geminates).

• Nam et al. (2009) model articulatory gestures as coupled oscillators. They assume that in-phase (0°) coupling is the most stable mode and anti-phase (180°) coupling the second most stable, as established in research on coordinated movement of human limbs; they also follow earlier work in associating the two modes with CV and VC sequences, respectively. They show that a model based on these two assumptions makes a number of correct predictions, including earlier acquisition of CV relative to VC by children.

Models of this type are an improvement over case-by-case reasoning in that they attempt to model larger systems, and often offer at least an in-principle solution to the problem of conflicting measures of effort. For example, Kirchner (2001b) incorporates into his metric of force expended the dimensions of speed of articulator movement (greater speed requires more force), distance traveled (a greater distance requires more force), and precision (precise gestures require several counteracting forces to keep the relevant articulator at exactly the right position). All these factors are collapsed onto a single dimension so that the net effort expended can be determined.

However, abstract models like these still require the researcher to choose some parameter of the model as the one that best represents ‘effort’; although the
methods described above involve reasonable choices, there is no guarantee that they are the best choices in terms of describing which sounds are more effortful than others. In addition, since models are by definition simplifications of real-world phenomena, there is always the danger of putting into these models what we expect to get out of them. It is not surprising, for example, that if we describe the articulatory apparatus as a means of applying forces to various masses, then it takes more force to cause a given mass to move farther or faster; or that if we assume that children are biased towards more stable CV sequences over less stable VC sequences, then they acquire CV sequences first.

In Experiment 1, I take a different approach altogether. Rather than trying to reason through whether lenited productions are more or less effortful than unlenited productions, I attempt to create conditions in the laboratory that will encourage subjects to use less articulatory effort than they otherwise would. If subjects favor one type of production in the ‘low-effort’ condition more than they do in the control condition, then we have evidence that that production requires less effort than productions that are not favored in the ‘low-effort’ condition. I am aware of only one other study along these lines in the literature: Walter (2008) reports experimental results showing that qualitative lenition is more likely in consonants with an identical consonant in an adjacent syllable.

In this study, I encourage effort reduction in subjects with intoxication. If this method is successful, it will tell us which segments are easier than others, but not why they are easier. As detailed in §3.2, intoxication impairs subject performance in a number of different ways, any of which might or might not be responsible for any effort reduction revealed by Experiment 1. Determining what factors contribute to articulatory effort and how they interact is an important topic for
future research. However, the necessary first step is to acquire more solid evidence as to which productions require more or less effort in the first place, and that is the goal of Experiment 1.

### 3.2 Physical and Linguistic Effects of Alcohol Consumption

Alcohol is a depressant and is known to impair cognitive and motor function (Chin and Pisoni 1997). At small doses, alcohol may actually increase performance slightly on some tasks (Chin and Pisoni 1997, 19-20,22); however, this effect seems to be limited to blood alcohol concentrations well below the level of .10 achieved in this study (Hollien and Martin 1996, 109-111). In general, intoxication impairs both speed and accuracy on tasks, and at levels above .10 impairs fine motor performance (Chin and Pisoni 1997, 21-22).

By far the most common documented effect of intoxication on speech is an overall lengthening effect; see, e.g., Lester and Skousen (1974, 233-234), Pisoni et al. (1986, 138), Johnson et al. (1990), Künzel (1992, 33,36), and Hollien and Martin (1996, 125). Intoxication has also been reported to result in expanded (Künzel 1992; Watanabe et al. 1994, 341) and more variable (Pisoni et al. 1986, 141) pitch ranges. Alcohol induces production errors at the segmental level (Künzel 1992; Hollien and Martin 1996, 125); specific errors that have been documented include lengthening of vowels and consonants, deletion, changes in nasality, and distortions of [s] (Lester and Skousen 1974; Künzel 1992; Pisoni et al. 1986). Künzel (1992, 33) observes substitution of [ʊ] for [d] and attributes the change to incom-
plete articulation. Purnell (2010) observes what appears to be a rotation of the vowel space (at least for front vowels) in intoxicated subjects.

There is some evidence that intoxication encourages devoicing, especially word-finally (Lester and Skousen 1974, 234). Pisoni et al. (1986) observe more lengthening in intoxicated speech in voiceless segments than in voiced segments. By contrast, Swartz (1992) observes no overall change in voice onset time in intoxicated speech. Watanabe et al. (1994, 346) document swelling of the vocal folds following consumption of alcohol, which may contribute to changes in subjects’ propensity to produce voicing.

In addition to its topical effects on the vocal folds, there are a number of other mechanisms by which alcohol may affect speech. Alcohol interferes with proprioeception (Wang et al. 1993; Tiplady et al. 2005), such that subjects become less accurate in their ability to move their arms a prespecified distance or draw figures of a prespecified size (when unable to see their arms or hands). Tiplady et al. also observe an increase in handwriting size under intoxication. The strongest effect is found for writing of unfamiliar orthographic characters, and the weakest effect for signatures, suggesting that alcohol has the least influence on highly practiced motor routines. Hellekant (1965) observes paralysis of certain fibers in the tongues of cats when exposed to alcohol; however, it is worth noting that Hellekant finds this effect only for alcohol concentrations above about 4.1 M, while the alcoholic drinks administered in this experiment (equal parts orange juice and 80-proof vodka) had a slightly lower concentration, approximately 3.4 M. Finally, in connection with the effects of intoxication on voicing in their study, Pisoni et al. (1986, 144) suggest that alcohol impairs subjects’ ability to coordinate different gestures, particularly oral gestures and control of the vocal folds.
Thus, alcohol affects subject performance in a number of ways, and it also affects speech. Several of the specific effects of alcohol parallel the metrics of articulatory effort discussed above:

- Reduced speed of subject performance may decrease the speed at which subjects can move their articulators
- Impaired accuracy may decrease subjects’ ability to effect precise placement of their articulators
- Alcohol as a depressant may limit the overall amount of energy subjects are able or willing to expend
- By impairing subjects’ overall cognitive abilities, alcohol may impair subjects’ ability to manage complex patterns of gestural coordination and cause them to revert to more ‘stable’ patterns

The fact that alcohol impairs cognitive and motor function in these ways, and the fact that intoxicated subjects do make more speech errors, suggest that alcohol consumption is a promising way to encourage subjects to expend less articulatory effort. Indeed, some of the speech errors documented under intoxication (such as substitution of [ð] for [d]) look tantalizingly like the types of lenition processes being investigated here.

However, intoxication is most relevant to the study of articulatory effort in general to the extent that it induces or exaggerates behavior that has other causes as well. In other words, if intoxication has some effect on speech production that causes subjects to lenite, but only intoxication affects speech in that way, then lenition by intoxicated subjects probably does not tell us much about what
drives lenition in general. It is unlikely that lenition never occurs in teetotalling communities! For some of the effects of alcohol (such as a general impairment of fine motor control), it seems quite likely that there are other ways to produce similar effects; for others, however (such as swelling of the vocal folds), it is less clear that we are not dealing with something alcohol-specific.

It is entirely possible, therefore, that in the final analysis intoxicated speech tells us little that is relevant to the role of articulatory effort in phonological patterns. But in the absence of more direct evidence concerning which productions are more difficult than others, I submit that this method is worth serious consideration. In addition, because the physical effects of alcohol are so varied, and because the articulatory apparatus is so complex, I am not comfortable pointing to any one effect of alcohol and claiming that that effect is the sole or primary cause of the results seen in Experiment 1. Indeed, it would be irresponsible to rule out intoxication entirely as a tool for investigating effort reduction simply because it might have other effects as well. No method of studying articulatory effort is perfect. We simply do not know the extent to which the results of this method are or are not specific to intoxicated speech – and we cannot know until we have tried it. I view this experimental method as an important early step in a research program that attempts to observe effort reduction in action.
3.3 Experiment 1

3.3.1 Design

Stimuli

Three sets of stimuli were used in the experiment; all were real words of English. A full list of stimuli is included in appendix 7.2. The ‘effective contrast’ (frequency-weighted neighborhood density, Ussishkin and Wedel (2009)) of each stimulus was calculated from the frequency data in the CELEX database (Baayen et al. 1995); within each set, stimuli were chosen from as narrow a range of EC values as possible to ensure comparability among different subsets of stimuli.

The ‘nasal-stop’ set consisted of 56 disyllabic words with an intervocalic nasal-stop cluster (e.g., amber); each possible stop ([b], [p], [d], [t], [g], or [k]) was represented by 10 stimuli, except for [g], for which there were 6 stimuli. All words had primary stress on the first syllable and no stress on the second syllable.

The ‘lenition’ set consisted of 72 disyllabic words with a single intervocalic stop (e.g., buggy); each possible stop was represented by 12 stimuli. All words had primary stress on the first syllable and no stress on the second syllable.

The ‘CVC’ set consisted of 136 words with the shape CVC; each of the vowels and diphthongs of English ([i], [ɪ], [ɛ], [æ], [ɑ], [ʌ], [u], [uː], [eɪ], [ou], [au], [aʊ], or [ɔi]) was represented by 12 stimuli, except for [v], [au], and [ɔi], for which there were fewer stimuli. All of the consonants in these stimuli were obstruents. These stimuli are treated as fillers in the analyses below and are not analyzed further.
Participants

Eight subjects participated in Experiment 1; all were undergraduate or graduate students at UC Santa Cruz and were naïve to the purposes of the experiment. One of these subjects, subject 00, was a graduate student in linguistics; her results do not differ noticeably from those of the other subjects and are therefore included. To be eligible to participate in the experiment, each subject was required to

- be 21 years old or older,
- regularly consume at least one alcoholic drink per week,
- have consumed enough alcohol in a single sitting within the previous year to have likely raised his or her blood alcohol content (BAC) to .10 or higher,
- not be at risk for alcoholism according to the Michigan Alcoholism Screening Test,
- not be taking any medications that proscribe the consumption of alcohol,
- not be pregnant,
- be a native speaker of English, and
- never have been diagnosed with any speech impediment or hearing deficit.

Table 3.1 gives each participant’s sex, whether that participant completed the sober or intoxicated condition first, and the participant’s BAC at the beginning and end of the recording session in the intoxicated condition.
Table 3.1: Participants in Experiment 1. BAC$_i$ and BAC$_f$ are the participant’s BAC at the beginning and end, respectively, of the recording session in the intoxicated condition.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>First Session</th>
<th>BAC$_i$</th>
<th>BAC$_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>F</td>
<td>sober</td>
<td>.10</td>
<td>.07</td>
</tr>
<tr>
<td>01</td>
<td>F</td>
<td>intoxicated</td>
<td>.09</td>
<td>.07</td>
</tr>
<tr>
<td>02</td>
<td>F</td>
<td>intoxicated</td>
<td>.11</td>
<td>.09</td>
</tr>
<tr>
<td>03</td>
<td>M</td>
<td>sober</td>
<td>.12</td>
<td>.10</td>
</tr>
<tr>
<td>04</td>
<td>M</td>
<td>sober</td>
<td>.10</td>
<td>.13</td>
</tr>
<tr>
<td>05</td>
<td>M</td>
<td>intoxicated</td>
<td>.10</td>
<td>.09</td>
</tr>
<tr>
<td>06</td>
<td>F</td>
<td>sober</td>
<td>.10</td>
<td>.07</td>
</tr>
<tr>
<td>07</td>
<td>M</td>
<td>sober</td>
<td>.10</td>
<td>.18</td>
</tr>
</tbody>
</table>

Procedure

Each subject was recorded in both the sober condition and the intoxicated condition. With the exception of subject 00, the two recording sessions took place on separate days; the order of the two sessions was varied across subjects. Sessions followed the following procedure:

1. For the intoxicated condition, the subject was asked not to consume alcohol for 12 hours before the session began, and not to eat for 2 hours before the session.

2. At the beginning of each session, the subject read out loud a printed list of the stimuli to ensure that the subject was familiar with all of the words and had seen them all recently.

3. In the intoxicated condition, alcohol was administered so that the subject had a BAC between .10 and .12. (This level is slightly above .08, the legal limit for driving in the United States.)
(a) The initial dose of alcohol (equal parts 80-proof vodka and orange juice) was half of what was calculated to be necessary to bring the subject’s BAC to between .10 and .12, based on the subject’s sex and weight.

(b) After the subject finished the initial dose, he or she was given a Breathalyzer test and, if necessary, received further smaller doses of alcohol until the window of .10 – .12 was reached.

4. The subject made the recording alone in a sound-attenuated booth. Stimuli were presented one by one in random order on a computer screen; the subject read each word in the frame sentence “I SAID __ already” and pressed the space bar to move on to the next word. Each stimulus was presented twice; the entire recording session took 20 – 30 minutes.

5. In the intoxicated condition, the subject remained in the lab until his or her BAC had declined to .04 or below.

3.3.2 Measurements

Subjects’ recordings were analyzed in Praat (Boersma and Weenink 2007). The following sections describe the landmarks that were identified for each type of stimulus and how they were placed.

All Stimuli

The following landmarks were identified in all tokens, regardless of stimulus type.

Beginning of utterance \( U_i \)  The beginning of the utterance (that is, the frame
sentence) was defined as the beginning of the (regular or irregular) glottal pulses of the I.

**End of utterance** ($U_f$) The end of the utterance was defined as the first point in the [i] of *already* that was followed by a period of silence (that is, a period with no energy at any point in the spectrum). Subjects commonly produced breathy voice at the end of the frame sentence, sometimes making such a point difficult to identify.

**Beginning of stimulus** ($W_i$) The way in which the beginning of the stimulus was defined depended on the initial segment of the stimulus.

For stimuli beginning with a vowel or sonorant consonant, $W_i$ was placed after the release of the [d] of *said* and any associated aspiration or other turbulent noise. If the subject paused between *said* and the stimulus, $W_i$ was placed at the first glottal pulse of the stimulus.

For stimuli beginning with a stop or affricate, $W_i$ was placed at the closure of the [d] of *said*. This location was chosen because in the vast majority of cases, subjects did not produce a separate release for the [d] of *said*, and it was therefore impossible to distinguish between the closure of the [d] and the closure of the initial consonant of the stimulus. In the few tokens that did have a separate release after *said*, $W_i$ was nevertheless placed at the closure of the [d] for the sake of consistency.

For stimuli beginning with a fricative, $W_i$ was placed at the beginning of the turbulent noise associated with the frication.

**End of stimulus** ($W_f$) The way in which the end of the stimulus was defined
depended on the final segment of the stimulus.

For stimuli ending with a stop, $W_f$ was placed after the release of the stop and any release or aspiration associated with it.

For stimuli ending with a fricative or affricate, $W_f$ was placed at the end of the turbulent noise associated with the frication.

For stimuli ending with a nasal, $W_f$ was placed at the point where the antiformants and decreased intensity associated with the nasal ended.

For stimuli ending with a vowel or sonorant consonant, $W_f$ was placed at the beginning of the formant transitions between the final segment of the stimulus and the [a] of already. In most cases, this final segment was either [i], in which case the beginning of the F2 transition was used, or [æ], in which case the beginning of the F3 transition was used.

**Closure of [d] of said ($S_{clo}$)** For all stimuli that did not begin with a stop or affricate (that is, those for which $W_i$ was not placed at the closure of the [d] of said), an additional landmark $S_{clo}$ was placed at the closure.

**End of voicing of [d] of said ($S_{voi}$)** The end of voicing of the [d] of said was defined as the last glottal pulse following $S_{clo}$. If there was continuous voicing from the release of the [d] through the beginning of the stimulus, $S_{voi}$ was not marked.

**Nasal-Stop Stimuli**

Figure 3.1 illustrates the additional landmarks that were identified in nasal-stop tokens.
Figure 3.1: Production of *anchor* by subject 01 in the sober condition

End of vowel (*A*) The end of the vowel was defined as the beginning of visible antiformants associated with the following nasal. If no antiformants were visible, *A* was placed at the sharp drop in intensity associated with the closure of the stop.

End of nasal (*B*) The end of the nasal was defined as the sharp drop in intensity associated with the closure of the stop. *B* was marked only if it was clearly distinct from both *A* and *C*, and if there were no frequencies above 1700 Hz immediately before the closure.

End of voicing (*C*) The end of voicing was defined as the end of the visible voice bar. In ambiguous cases, I relied on the glottal pulses identified by Praat for guidance.
Figure 3.2: Production of labor by subject 01 in the sober condition

Release of consonant ($D$) The release of the consonant was defined as the vertical dark bar after the closure. If there was no clear release, $D$ was not marked.

Lenition Stimuli

Figure 3.2 illustrates the additional landmarks that were identified in lenition tokens.

Beginning of consonant ($E$) The beginning of the consonant was defined as the sharp drop in intensity following the preceding vowel.

End of voicing ($F$) As in the nasal-stop stimuli, the end of voicing was defined as the end of the visible voice bar.
Release of consonant \((G)\) As in the nasal-stop stimuli, the release of the consonant was defined as the vertical dark bar after the closure.

End of consonant \((H)\) The end of the consonant was defined as the first regular glottal striation in the following vowel.

3.3.3 Results

All Stimuli

Duration of frame sentence \((U_f - U_i)\). Given the common finding (noted above) that intoxicated speech is slower than sober speech, examining the effect of intoxication on the duration of the entire frame sentence allows us to establish a baseline for comparison with other duration measurements that are more directly relevant to lenition. Recorded tokens were not included from this analysis if they were clipped at either end (because the subject began speaking too early or pressed the space bar before finishing the frame sentence) or if the subject hesitated noticeably during the sentence.

Figure 3.3 shows results for the duration of the frame sentence. Each sub-graph shows results for one subject; subjects 05 and 07 are omitted because too many recordings were clipped to yield a reasonable sample size. Each point represents a single stimulus word; the x-axis shows the duration of the frame sentence for that word in the sober condition, and the y-axis shows the duration in the intoxicated condition, each averaged over two repetitions.

The dotted line in each graph is the line \(x = y\); if alcohol has no effect on the duration of the frame sentence, then the data points should lie along this line. The solid lines represent a linear mixed-effects model predicting intoxicated
Figure 3.3: Duration of frame sentence by subject for nasal-stop and lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

duration from sober duration with by-subject slopes and intercepts. Outliers – defined as data points with residuals more than 2.5 standard deviations from the mean (Baayen 2008, 188-192) – are omitted from the final models and plotted in light gray. A star next to a subject number means that for that subject, the slope of the regression line is significantly greater than 0 and less than 1. A star at the top or bottom of a subject’s graph means that for that subject, a paired t-test reveals significantly longer or shorter durations, respectively, in the intoxicated condition. For both tests, the cutoff for significance is $\alpha = .05$. All subsequent graphs in this section have the same layout.
These results suggest the expected overall pattern of lengthening in the intoxicated condition; subjects 00, 01, 02, and 04 exhibit significantly longer durations in the intoxicated condition. (However, subject 03 shows the opposite pattern, and subject 06 shows no effect.) Note, though, that no subject exhibits uniform lengthening in the intoxicated condition. Except for subject 06, all of the subjects display what I will refer to as the ‘X-pattern’: the slope of the regression line is shallower than 1, and the regression line crosses the line $x = y$ near or within the main grouping of data points. Thus, the ‘X-pattern’ reveals a tendency whereby, in the intoxicated condition, relatively more time is added to (or removed from) utterances with extreme duration values (that is, very long or very short utterances) than for utterances with less extreme values. When the regression line crosses $x = y$ in the middle of the data, subjects exhibit this avoidance of extreme values at both ends of the scale: very long utterances are shortened, and very short utterances are lengthened. Figure 3.4 schematizes how this compression leads to the ‘X-pattern’. In other words, subjects appear to exhibit less variance in utterance duration in the intoxicated condition.
Figure 3.5: Duration of stimulus by subject for nasal-stop and lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

**Duration of target word** ($W_f - W_i$). Like the duration of the frame sentence, the duration of the target word is relevant to determining whether there is an overall lengthening effect in the intoxicated condition.

Figure 3.5 shows results for the duration of the stimulus word. This measure shows some lengthening in the intoxicated condition relative to the sober condition: duration increases for subjects 01, 02, 04, 05, and 07, but decreases for subjects 03 and 06.

**Length of voicing of [d] of said** ($S_{\text{voi}} - S_{\text{clo}}$). Word-final devoicing of voiced obstruents is a typologically common process that has been argued to involve
Figure 3.6: Duration of voicing of [d] of *said* by subject for nasal-stop and lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

Voicing duration in "said", sober (s)
Voicing duration in "said", intoxicated (s)

effort reduction (see Westbury and Keating (1986, 156-157), among others). If intoxication encourages final devoicing, then the final voiced [d] of *said* in the frame sentence should have less voicing in the intoxicated condition than in the sober condition. In addition, examining the voicing of the [d] of *said* allows us to distinguish between any effects that intoxication has on postnasal and intervocalic stops from the effects that it has on stops generally. Voicing of the [d] is analyzed only when the target word begins with a voiceless stop or sibilant: before voiced obstruents and sonorants, there is the possibility that the [d] may assimilate in voicing to the following segment, resulting in an *increase* in voicing; before the
non-sibilant voiceless fricatives [f], [θ], and [h], the voicing of [d] often continued well into the following segment.

Figure 3.6 shows results for the duration of voicing in the [d] of *said*. The plotting symbol for each point is the initial consonant of the stimulus word; ⟨c⟩ represents [f], and ⟨S⟩ represents [ʃ]. There is no support for an overall pattern of devoicing: subject 03 has decreased voicing in the intoxicated condition, while subjects 05 and 06 have increased voicing. Although several subjects exhibit the ‘X-pattern’ (00, 01, 02, 05, 06, and 07), two have no significant correlation between voicing durations in the two conditions (03 and 04). This weak result is not surprising, given the fact that utterances are paired across conditions by the stimulus word. We expect the duration of the word to be highly dependent on the identity of the word, and, thus, we should find a strong correlation between the duration of the word in the sober and intoxicated conditions. However, the identity of the word is likely to have a smaller effect on the voicing of the final consonant of the word that immediately precedes it; thus, the correlations between the sober and intoxicated conditions are weaker.

**Nasal-Stop Stimuli**

**Duration of consonant.** Voiceless obstruents tend to be longer than voiced obstruents (Kingston and Diehl 1994, 441). If intoxication encourages postnasal voicing, then consonants should be shorter in the intoxicated condition. Consonant duration was calculated in two ways: from the end of the vowel ($D - A$), and from the end of the nasal, if present ($D - B$).

Figure 3.7 shows results for the duration of the consonant, as measured from the end of the vowel (top graph) and from the end of the nasal (bottom graph).
Figure 3.7: Duration of consonant from end of vowel (top) and nasal (bottom) by subject for nasal-stop stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.
The plotting symbol is the stop in the nasal-stop sequence of the stimulus word. As was the case for the two global duration measures considered above, there appears to be a general trend of lengthening in the intoxicated condition; only subjects 00, 04, and 05 fail to exhibit a significant increase in consonant duration for at least one measurement. The ‘X-pattern’, where the regression lines have a slope shallower than 1, is consistent across subjects and measurements.

This lengthening effect is not consistent with postnasal voicing. However, it is possible that the lengthening seen here is not due to postnasal devoicing in the intoxicated condition, but rather to the overall lengthening of intoxicated speech.

**Duration of voicing.** If intoxication encourages postnasal voicing, then voicing should increase in the intoxicated condition. As for overall duration, voicing duration was calculated in two ways: from the end of the vowel \( (C - A) \), and from the end of the nasal, if present \( (C - B) \). In each case, if the consonant was fully voiced, then voicing duration was measured up to the release of the consonant \( (D) \).

Figure 3.8 shows results for the duration of voicing. For both measures, four subjects exhibit significantly increased voicing. When voicing is measured from the end of the nasal, the remaining three subjects exhibit decreased voicing; when voicing is measured from the end of the vowel, the remaining three subjects show no effect. Because a large number of recorded tokens contained fully voiced consonants (as shown by the clustering of data points for voicing proportion in figure 3.9 around 1), it is likely that these results are parasitic on the effect of intoxication on overall consonant length. Note that in most cases, a subject who shows a significant increase or decrease in voicing duration in the intoxicated condition for some measure also shows an increase or decrease in consonant duration for the
Figure 3.8: Duration of voicing from end of vowel (top) and nasal (bottom) by subject for nasal-stop stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.
corresponding measure. In addition, no subject shows a significant effect in one direction for consonant duration and in another direction for voicing duration.

**Proportion of consonant that is voiced.** Since intoxication appears to have a slight lengthening effect on consonant duration, any increase in voicing duration may be due solely to the change in consonant length, and not to a change in subjects’ propensity to produce voicing. Measuring voicing in terms of the proportion of the consonant that is voiced is one way to factor out the effect of overall consonant length. As above, voicing proportion was calculated in two ways: from the end of the vowel \( \frac{C-A}{D-A} \), and from the end of the nasal, if present \( \frac{C-B}{D-B} \).

Figure 3.9 shows results for voicing proportion. Although linear models were fitted to these results as to the others, they are not entirely appropriate since the dependent variable, a proportion, is bounded by 0 and 1; thus, the ‘X-pattern’ exhibited for voicing measured from the end of the nasal may not be reliable. Only a few subjects show a consistent change in voicing proportion in the intoxicated condition, and the direction of the change is not consistent (an increase in some cases, a decrease in others). This measure, therefore, does not provide evidence that there was an overall change in the amount of voicing in the intoxicated condition.

**Lenition Stimuli**

**Duration of consonant** \((G – E)\). If intoxication encourages intervocalic voicing, then consonants should be shorter in the intoxicated condition. Duration was calculated to the release of the consonant, as with the nasal-stop stimuli.

Figure 3.10 shows results for consonant duration. There appears to be a
Figure 3.9: Proportion of closure that is voiced from end of vowel (top) and nasal (bottom) by subject for nasal-stop stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.
Figure 3.10: Duration of consonant by subject for lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$; stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

lengthening effect; five subjects (02, 03, 05, 06, and 07) show significantly increased duration in the intoxicated condition, and no subject has the opposite effect. Stronger and more consistent, though, is the ‘X-pattern’ noted above: the regression line for every subject has a slope between 0 and 1 (and significantly different from either).

**Duration of voicing** $(F - E)$. If intoxication encourages intervocalic voicing, then voicing should increase in the intoxicated condition. As for nasal-stop stimuli, if the consonant was fully voiced, voicing was measured to the release $(G)$.

Figure 3.11 shows results for voicing duration. There appears to be no overall shift in voicing duration in the intoxicated condition; only two subjects have a
Figure 3.11: Duration of voicing by subject for lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$; stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

significant effect (in different directions). However, every subject except subject 03 exhibits a significant ‘X-pattern’: highly voiced consonants become slightly less voiced in the intoxicated condition, and consonants with very little voicing acquire slightly more. Subjects in the intoxicated condition appear to avoid extreme values of voicing at either end of the spectrum, and instead favor intermediate voicing durations.

Proportion of consonant that is voiced ($\frac{F - E}{G - E}$). As with the nasal-stop stimuli, voicing during the consonant was also calculated in terms of the proportion of the consonant that is voiced, in order to factor out the effect of changes in overall consonant length.
Figure 3.12: Proportion of closure that is voiced by subject for lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = 0.05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

Figure 3.12 shows results for voicing proportion. Only subject 03 exhibits a change in overall voicing in the intoxicated condition (a slight decrease); all other subjects except subject 05 have the ‘X-pattern’ but no general shift in the amount of voicing during the consonant. Thus, this measure does not support a pattern of intervocalic voicing.

Duration of burst ($H - G$). The duration of the consonant’s burst and any subsequent aspiration (referred to here simply as “burst duration”) is longer for voiceless stops than for voiced stops. (In English, the extra duration for voiceless stops is due at least in part to the aspiration that follows them, even in intervocalic
Figure 3.13: Duration of burst by subject for lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom duration is significantly greater or less, respectively, in the intoxicated condition.

Figure 3.13 shows results for burst duration. Four subjects show increased burst duration in the intoxicated condition (00, 01, 03, and 06); five show the ‘X-pattern’ (02, 04, 05, 06, and 07). Burst duration does not support a pattern of intervocalic voicing in the intoxicated condition; if anything, it suggests a change in the opposite direction.

Ratio of minimum intensity in consonant to maximum intensity in
Figure 3.14: Ratio of intensity of consonant to preceding vowel (top) and following vowel (bottom) by subject for lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at $\alpha = .05$. Stars at the top or bottom of the graphs mark subjects for whom the ratio is significantly greater or less, respectively, in the intoxicated condition.
vowel. Voiced stops are expected to be more intense than voiceless stops during closure because of the energy added by the voicing component; spirants are expected to be more intense than stops because of the energy added by the turbulent airflow of frication. If intoxication encourages intervocalic voicing or spirantization, the intensity of the consonant should increase in the intoxicated condition. With this measure, the intensity of the consonant (represented by the minimum intensity during the consonant closure) is relativized to the maximum intensity of the preceding or following vowel. For this measure, and for all subsequent measures involving intensity, subject 04 is omitted: this subject’s recordings had significant clipping in the sober condition.

Figure 3.14 shows results for intensity ratios relative to the preceding vowel (top graph) and following vowel (bottom graph). Although subject 00 has increased consonant intensity relative to the preceding vowel in the intoxicated condition, there is no other significant increase in consonant intensity, and there are several significant decreases. Thus, these measures provide no evidence for an overall pattern of intervocalic voicing or spirantization in the intoxicated condition. The ‘X-pattern’, on the other hand, is significant for every subject for both measures: in the intoxicated condition, subjects avoid both very intense consonants and very quiet ones.

Smallest/largest slope of intensity contour in consonant. Another measure of the intensity of the consonant relative to the surrounding vowels involves the slope of intensity contour (Kingston 2008). The quieter the consonant, the steeper the slope of the contour will be leading into and out of the consonant (because of the large transition between the loud vowel and the quiet consonant); conversely, the more intense the consonant, the shallower the slope will be. We
Figure 3.15: Intoxicated production of *rapid* (top) and sober production of *rabbit* (bottom) by subject 02, with intensity contours; arrows show maximum slope of intensity contour during consonant.

![Graph of “rapid”](image)

![Graph of “rabbit”](image)

...can get a global measure of the slope of the intensity contour leading into the consonant by finding its smallest (most negative) slope; similarly, we can measure the slope leading out of the consonant by finding its largest slope. (See figure 3.15 for an illustration.) If intoxication induces intervocalic voicing or spirantization, then each measure should be significantly closer to zero in the intoxicated condition.

Figure 3.16 shows results for minimum (top graph) and maximum (bottom graph) slopes of the intensity contour. There is no general pattern in terms of an
Figure 3.16: Minimum (top) and maximum (bottom) slope of intensity contour during consonant by subject for lenition stimuli. Stars next to subject numbers mark subjects with a slope significantly different from 0 and 1 at \( \alpha = .05 \). Stars at the top or bottom of the graphs mark subjects for whom the slope is significantly greater or less, respectively, in the intoxicated condition.
overall shift in the intoxicated condition: for each measure, some subjects have significant increases, while others have significant decreases. These measures, like most of those before them, do not support a pattern of intervocalic voicing or spirantization. The ‘X-pattern’, however, remains robust: with only two exceptions (subjects 03 and 05 for maximum slope), the regression lines uniformly have a slope between 0 and 1.

3.3.4 Discussion

Overall Results

Table 3.2 summarizes significant results for each of the measures discussed above. For each subject and each measure, there are two possible significant effects. The first is whether the subject displays the ‘X-pattern’ for that measure: that is, whether the slope of the regression line for that subject is significantly less than 1 (but greater than 0). Subjects with this pattern appear to avoid ‘extreme’ articulations at one or both ends of the relevant scale, and to favor instead productions with intermediate values. For example, for subject 06, if a consonant in one of the lenition stimuli had a great deal of voicing in the sober condition as measured by proportion of the closure that was voiced, then that consonant in the same stimulus in the intoxicated condition tended to have slightly less voicing. But if a consonant had very little voicing in the sober condition, then it tended to have slightly more in the intoxicated condition. Significant ‘X-patterns’ are shown with gray cells.

The second type of significant effect is an overall difference in the measured quantity between the two conditions, as determined by a paired t-test. Subjects
Table 3.2: Summary of significant effects. A gray cell denotes a regression line with a slope significantly different from 0 and 1. An arrow denotes a significant difference between the sober and intoxicated conditions as determined by a paired t-test. The direction of the arrow shows whether the value increased or decreased in the intoxicated condition; a double arrow shows an effect in the expected direction, a single arrow the opposite direction.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>00 01 02 03 04 05 06 07</td>
</tr>
<tr>
<td><strong>All Stimuli</strong></td>
<td></td>
</tr>
<tr>
<td>Dur. frame sentence</td>
<td>↑↑↑↑↓↑</td>
</tr>
<tr>
<td>Dur. stimulus</td>
<td>↑↑↑↓↑↑↑</td>
</tr>
<tr>
<td>Dur. voicing in <em>sai</em>[d]</td>
<td>↓↑↑↑</td>
</tr>
<tr>
<td><strong>Nasal-Stop Stimuli</strong></td>
<td></td>
</tr>
<tr>
<td>Dur. consonant (from V)</td>
<td>↑↑↑↑</td>
</tr>
<tr>
<td>Dur. consonant (from N)</td>
<td>↓↑↑↑</td>
</tr>
<tr>
<td>Dur. voicing (from V)</td>
<td>↓↑↑↑↑</td>
</tr>
<tr>
<td>Dur. voicing (from N)</td>
<td>↓↑↑↑↑↑↑</td>
</tr>
<tr>
<td>Prop. closure voiced (from V)</td>
<td>↑↓↓↑</td>
</tr>
<tr>
<td>Prop. closure voiced (from N)</td>
<td>↓↓↑↑</td>
</tr>
<tr>
<td><strong>Lenition Stimuli</strong></td>
<td></td>
</tr>
<tr>
<td>Dur. consonant</td>
<td>↑↑↑↑↑↑</td>
</tr>
<tr>
<td>Dur. voicing</td>
<td>↓↑↓</td>
</tr>
<tr>
<td>Prop. closure voiced</td>
<td>↓↑</td>
</tr>
<tr>
<td>Dur. burst</td>
<td>↑↑↑↑</td>
</tr>
<tr>
<td>Int. ratio, C/V1</td>
<td>↑↑↑↓↓</td>
</tr>
<tr>
<td>Int. ratio, C/V2</td>
<td>↓↓</td>
</tr>
<tr>
<td>Min. slope, int. contour</td>
<td>↑↑↑↓</td>
</tr>
<tr>
<td>Max. slope, int. contour</td>
<td>↓↑↑</td>
</tr>
</tbody>
</table>

63
with this pattern display an overall preference for one end of the scale over the other in the intoxicated condition. For example, subject 03 had more intense consonants in the intoxicated condition, as measured by the maximum slope of the intensity contour during the consonant, than in the sober condition. Significant differences of this type are shown with an arrow. The direction of the arrow shows whether the measured value increased (up) or decreased (down) in the intoxicated condition. The shape of the arrow shows whether the change was in the expected direction (that is, the direction corresponding to intervocalic voicing/spirantization, postnasal voicing, or final devoicing): double arrows show expected changes, and single arrows unexpected ones. For the overall duration of the frame sentence and the stimulus word, neither change is necessarily predicted by the sound patterns under consideration; for these measures, the ‘expected’ change (as marked in the table) is an increase in duration, given previous findings in the literature that intoxication leads to slower speech.

Note that these two patterns are not mutually exclusive: it is possible for a subject to display the ‘X-pattern’ in addition to an overall preference for one end of the scale over the other for the same measure. For example, subject 06 shows longer consonants (measured from the end of the nasal) in the nasal-stop stimuli in the intoxicated condition. However, the effect is strongest for those stimuli whose consonants are very short in the sober condition; for stimuli whose consonants are already very long in the sober condition, there is little or no increase in the intoxicated condition. In other words, the range of consonant duration that subject 06 is willing to produce is smaller in the intoxicated condition than in the sober condition, even though the midpoint of that range is relatively high compared to the range in the sober condition.
Table 3.2 provides little, if any, support for the prediction that intoxicated speech mirrors attested phonological patterns. Significant changes in the intoxicated condition that correspond to the attested patterns of intervocalic voicing and spirantization, postnasal voicing, or final devoicing are few and far between, and a change in the expected direction for one subject is almost always accompanied by a change in the opposite direction for that measure for some other subject. This is the case even if we disregard all of the measures based on duration, where the general lengthening of intoxicated speech may be a confounding factor.

The pattern for which table 3.2 provides strong support is the ‘X-pattern’: when intoxicated, subjects are less willing to make ‘extreme’ productions. There are only a handful of cases where this pattern is not significant. Thus, the articulatory space of intoxicated subjects appears to be compressed; rather than (or in addition to) favoring one end of the scale over the other, subjects are reverting to gestures in the middle of the articulatory space.

The recurrent ‘X-pattern’ is consistent with the idea that intoxicated subjects expend less effort. If extreme gestures require more effort to produce than gestures in the middle of the articulatory space, then the compression of the articulatory space observed in the intoxicated condition means that subjects expended less effort after drinking alcohol than they did while they were sober. As discussed in chapter 5, there are in fact a number of proposals in the phonological literature linking articulatory effort with the extremes of the articulatory space; thus, the results of this experiment speak directly to particular theories of the relationship between articulatory effort and phonological patterns.
Effect of BAC

If the primary effect of alcohol on production is to cause contraction of the articulatory space, then we might expect higher levels of intoxication to cause more contraction. Because BAC varied both within and across subjects during Experiment 1 (as shown in table 3.1), it is possible the data could provide us with evidence for such an effect.

I used the within-subject variation in BAC to determine whether there was a within-subject effect of BAC on the degree of articulatory contraction. Each recording session was divided into two parts: the first half of the trials (containing the first repetition of each stimulus word) and the second half (containing the second repetition). The initial and final BAC measures for each subject were used as a rough approximation of that subject’s BAC during the first and second halves of the intoxicated recording session, respectively.

As above, for each measure of lenition, I built a linear mixed-effects model predicting values in the intoxicated condition from values in the sober condition. Productions from the first half of the intoxicated recording session for each subject were paired with productions from the first half of the sober recording session for the same subject; productions from the second half of each session were paired in the same way. Each model includes random slopes and intercepts for each half of the recording session for each subject. If higher BACs do indeed cause greater contraction of the articulatory space, then the slope of the regression line should be systematically smaller for the half of the recording session in which a given subject had a higher BAC, and larger for the other half of the recording session for the same subject.
As it turns out, there is no systematic difference in slope between subjects’ productions with higher or lower BACs. A paired t-test of the slope of the regression line for all measures and subjects reveals no significant difference between more and less intoxicated parts of the recording session ($p = .71$). However, there are several reasons not to be surprised at this null result:

- The within-subject differences in BAC are very small (usually between .01 and .03).

- For most subjects, BAC was higher at the beginning of the recording session; for subjects 04 and 07, BAC was higher at the end. It is possible that there are independent effects of timing within each session that are masking any effects of relative BAC.

- Although we know each subject’s BAC at the beginning and end of the experiment, we do not know what trajectory the subject’s BAC took *during* the recording session. For example, subject 00 began the intoxicated recording session with a BAC of .10 and ended it with a BAC of .07. However, we do not know whether the subject peaked at .10 and experienced a monotonic decline in BAC during the session, or whether the subject peaked later during the session and only then experienced a drop in BAC. Thus, beginning and ending BAC give us only a rough approximation of subjects’ BAC throughout the recording session.

- For subjects 04 and 07, who had a higher BAC at the end of the session than at the beginning, we do not know whether they peaked during or after the recording session – that is, whether the value measured at the
end of the session was on the rising or falling side of the BAC curve. This difference is important because of the phenomenon known as the Mellanby effect, whereby perforance of intoxicated subjects is more impaired while BAC is rising than it is at same BAC when BAC is falling (Wang et al. 1993).

Thus, although a post-hoc analysis of the effects of variation in BAC in Experiment 1 provides no evidence that higher BACs lead to greater contraction of the articulatory space, it is possible that an experiment specifically designed to answer this question would find an effect.

### 3.4 Conclusion

Experiment 1 was designed to determine whether intoxication – hypothesized to induce subjects to expend less articulatory effort – results in productions that resemble certain phonological patterns that have been claimed to involve effort reduction, including final devoicing, postnasal voicing, and intervocalic voicing and spirantization. The results of the experiment show that subjects did not systematically alter their productions when intoxicated in a manner corresponding to these patterns; instead, the most robust effect observed was a compression of the articulatory space in the intoxicated condition. Although there is no guarantee that alcohol consumption does in fact encourage subjects to expend less articulatory effort, the articulatory compression that they exhibit is plausibly interpreted as effort reduction and should not be lightly dismissed. Chapter 5 considers the implications of these results both for phonological theories of articulatory effort and for our account of lenition patterns in particular.