Frame analysis and the field of linguistics

In mainstream Anglo-American linguistics of the 20th century, the use of the term ‘frame’ appears to have developed quite independently of its history in the social sciences, and came about largely due to the influential work of one scholar, Charles Fillmore. In the 1950s Fillmore was examining the principles behind the co-occurrences of strings of words, which led him to examine the distributional properties of individual verbs (Fillmore 1961) in what could be called syntactic frames. However Fillmore moved away from the dominant school of linguistics in which he developed these ideas, namely transformational generative grammar; he put more emphasis on the semantics of the verbs as a driving force behind their co-occurrence patterns with nouns, rather than focussing on what would be predicted by the syntactic rules of grammar from Noam Chomsky’s generative theories. Also diverging from the predominant idea in American linguistics of the 1970s that word meanings were best characterized in terms of sets of necessary and sufficient conditions, Fillmore (1975) presented a “scenes-and-frames paradigm,” extending the notion of frame from syntax to semantics. Here the term ‘frame’ was used as a way of talking about the linguistic choices (words, grammatical rules, or other linguistics categories) that can get associated with prototypical instances of scenes: standard action scenarios within a given culture. Fillmore developed the notion of frame semantics (1982) for some time, and his work laid the foundations for approaches to semantics and grammar that are now known under the collective label of ‘cognitive linguistics’ (Cienki 2008).

Since then, frame analysis has been approached from various directions within linguistics. It has focussed on the lexical level (e.g., Baker 2006; Goatly 2007) as well as the level of grammar (leading, for example, to the theory of construction grammar; Goldberg 1995). What these different types of research share is an implicit focus on the analysis of written texts, including the transcription of spoken discourse.

The bias in modern linguistics toward formal (Chomskian generative) approaches and the prioritization of the written sentence as the unit of analysis is partly responsible for the relegation of spoken language research to the field of anthropology (witness the development
of ‘the ethnography of speaking’ as a sub-field in anthropology) and to sociology where it gave rise to Conversation Analysis. However more functionally oriented approaches began to gain ground and gain recognition within the field of linguistics in the late 20th century. Combined with the rise of digital technology which has made the recording and analysis of audio and video more accessible to more people, there is now an increasing amount of research on spoken language data from various theoretical perspectives.

Along with this move there has been greater acknowledgement that spoken language is organized according to some very different principles than written discourse is. For example, there has been increased attention in linguistics to intonation units as a fundamental structure of spoken language (Chafe 1994; DuBois et al. 1993; Wennerstrom 2001). Intonation units not only represent a convergence of prosodic features, but also display particular syntactic patterns (Croft 1995; Du Bois 2003), and therefore present information fundamentally differently than written sentences do. Current research is also focussing on other embodied aspects of spoken communication, such as gesture (Cienki & Müller 2008; Kendon 2004; McNeill 1992) and eye gaze (e.g., Streeck 1993), which have been found to align with intonation units.

**Spoken language framing**

Taking the term ‘frame’ as a way of talking about linguistic choices, what I propose in this paper is a broadening of the scope of the term to account for certain choices, either conscious or unconscious, which are made when one is expressing him/herself when speaking. Gibson (1979) used the term ‘affordances’ to refer to the interactive possibilities of a particular object or environment. So, chairs afford sitting on, and handles afford grasping. We can say that writing and speaking afford expression in different ways due to the properties of the media they involve. The face-to-face situation of the prototypical conversation in person is what Clark (1973) refers to as ‘the canonical encounter’ for interaction between two people. In this situation, the two (sighted and hearing) individuals can both hear and see each other because of their positions facing and looking at each other. What I will call ‘spoken language framing’ concerns the choices of expression for talk which take advantage of these affordances of linguistic communication in, or as if in, the canonical encounter situation. Because the situation involves hearing tone of voice, intonation, pauses, and seeing facial expressions, body position, and manual gestures, the semantic information gleaned through the canonical encounter goes beyond the level of information given in the words used which can be represented as propositional arguments (à la Kintsch 1998), and includes notions normally considered part of the separate field of pragmatics, such as emotion, and the speaker’s perspective on his/her discourse. In addition, speech (and the maximizing of the features of one’s talk as talk via spoken language framing) is inherently dynamic, with change over time as one of its properties, making this process qualitatively different from a written document (a product of/for communication).

The categories of informal versus formal language should be mentioned here in connection with the contrast between spoken and written forms of language. It is worth noting that the two parameters clearly do not line up in parallel with each other; we have informal varieties of writing – such as most instant messaging (IM), and formal talk – as in official speeches. What is important here is that political debates are a mixed genre of talk which is more toward the formal end of the spectrum, in contrast to the ur-site of spoken language use, which is informal conversation. The claim here is that spoken language framing involves
using affordances particular to speech as opposed to writing, and therefore possibly using elements characteristics of informal, conversational interaction.

The units operationalizing spoken language framing that will be analyzed include the use of the following.

- **pronouns**

Spoken discourse is different from written discourse in terms of the frequency with which certain words are used, as studies of large databases (known as linguistic corpora) of written texts and transcribed speech have shown (Biber et al. 1999). One difference to be investigated here is the use of pronouns. In general, the corpus research has shown that pronouns are used more in spoken than in written discourse. Speech, and face-to-face conversation in particular, provides grounding for the use of pronouns, and especially for first-person (in English, *I, we*) and second-person (*you*) pronouns, words which are ‘shifters’ (Jakobson 1957/1990) whose meaning depends on the context of who is using them and to whom they are being addressed. First- and second-person pronouns therefore appear much more frequently in spoken language use, and the use of both types presumes a frame of some kind of interaction between speaker and hearer. Pronoun use in written texts is more commonly restricted, particularly in formal writing, to the third-person (*he, she, it, they*). The argument here is thus that a high level of pronoun use, especially of first- and second-person pronouns, is one characteristic of spoken language framing.

- **colloquial language (vocabulary and pronunciation)**

Two categories will be examined here which mark language use that is more typical of (spoken) conversation. One is colloquial vocabulary: word choice from a more informal register, e.g., *folks* as opposed to *people*. The second involves pronunciation, namely the partial versus full articulation of words, e.g., *goin’* versus *going*; the category of partial articulation will include blending of word forms, such as *gonna* instead of *going to*.

- **syntactic structures**

Given the static nature of written language, readers can take time to make sense out of complex phrases (e.g., *the previously occupied place was now free*) because all parts of the construction are available in front of them, and so they can go back and reread portions of text which they might not have understood immediately. In speech, such phrases are more commonly broken up into separate clauses (*the place that was occupied before was now free*). In writing or formal speaking, clauses are connected and links between ideas are made explicit, whereas in informal speech elements understood from the previous verbal context can be left out. Similarly, referents in written texts have to be made explicit for a reader to be able to track them, whereas in spoken language use, particularly if the referent is the speaker him/herself or your addressee, it will be clear from context and need not be mentioned (“Wanna do that?” “Mm, don’t think so.”). The preferred grammatical structure of transitive clauses (subject-verb-object) in conversation is for the subject to be a pronoun (or simply left out as understood from context), rather than a full noun phrase (Du Bois 2003). This aligns with the more frequent use of pronouns in spoken versus written discourse, mentioned above.

- **intonation patterns**

The speech in the text sample was broken up into intonation units. Whereas the field of conversation analysis uses the turn-at-talk as a basic unit of transcription for studying
interaction between participants, cognitive linguistic research on spoken language focusses more often on the level of the individual utterances as they reflect the speaker’s processes of formulating his/her own ideas. Intonation units have been found to correlate with the development of new ideas in conversation, such that each intonation unit represents an idea unit (Chafe 1994). Research on gesture with speech on processes of formulating one’s thoughts for speech (thinking-for-speaking) supports this claim (McNeill & Duncan 2000).

Patterns within each intonation unit were studied using a variation on the Tone and Break Indices (ToBI) system of transcribing intonation (http://anita.simmons.edu/~tobi/). The ToBI system provides a way to mark the prominent points of an intonation contour which are meaningful for speakers of the language, namely stressed syllables (ones with greater perceived volume) and points where pitch contrasts occur. Rather than focussing on the acoustic details of the entire speech stream, the prominent elements picked out with the ToBI transcription are particular to each language (even to different language varieties, like varieties of English), and so provide what is considered by many linguists to be a guide to the “phonology” of intonation for the particular language (variety). Details on the identification of intonation units and of ToBI transcription will be provided in the Analysis section below.

Co-verbal behavior: eye gaze

The direction of the speaker’s gaze plays a role that speaker and addressee are often not consciously concerned with, yet it is an important role in how the speaker is interacting with the addressee (Kleinke 1986). Speakers in the culture under investigation here (mainstream American) normally maintain off and on eye contact with their addressee, often looking away while formulating thoughts, and looking the addressee in the eye when completing utterance of the idea, often leading to the end of the speaker’s turn. The assumption for this study is that using this pattern with the TV camera as the addressee will invoke a more natural conversational framing by the speaker than if s/he looks off-camera consistently while speaking.

Manual gesture with speech is another obvious co-verbal behavior which could have been studied. However the speakers’ hands were not visible with sufficient consistency in the televised broadcast of the debate to allow for analysis. This “talking head” bias of the visual framing of speakers is unfortunately typical of contemporary television in the U.S. and many countries.

Wrap-up:

The argument put forth is that these aspects can play a role in framing with spoken language which are importantly different from those available in written language. The framing by a speaker can be more heavily skewed towards the sensory (audio-visual) representation of him/herself, more readily reflecting the speaker’s affective stance towards the ideas s/he is uttering. Since co-verbal features of talk can serve particular interactional functions, speakers can also use them to a greater or lesser degree to frame their talk as a conversational encounter with the audience. In the discussion section of the paper I will also consider how this employment of dialogic norms could facilitate taking part in ‘fictive interaction’ (Pascual 2008) between audience and speaker. Indeed, recent research in neuroscience points to how fuller exploitation of audio-visual means of communication can bring about greater empathy with the speaker by the listener/viewer. We will see how for certain approaches to party politics and especially to pre-election campaigning, this form of self-presentation can be particularly desirable.
The data analyzed:

I will present a comparison of the use of spoken language framing devices by the two most recent candidates for vice president in the US, Joseph Biden and Sarah Palin, in their October 2008 televised debate. This event was chosen due to the extensive media attention it received, giving it an influential role in the campaign. It was the one forum in which these two politicians were able to discuss the same issues, with equal time per speaker, thus also affording a comparison between them. The structure of the debate was such that each candidate had 90 seconds to respond to a direct question and then an additional two minutes for rebuttal and follow-up. The debate included approximately the same amount of words by each candidate: 7302 for Biden, and 7760 for Palin.

The complete video is available at [http://www.youtube.com/watch?v=89FbCPzAsRA](http://www.youtube.com/watch?v=89FbCPzAsRA). A transcript of the debate is available on the website of the US Commission on Presidential Debates: [http://www.debates.org/pages/trans2008b.html](http://www.debates.org/pages/trans2008b.html). For some analyses, portions of the debate were re-transcribed to make the transcript conform more closely to what was actually uttered. This is described in more detail below.

Analysis:

-use of pronouns

The contributions of Biden and Palin in the debate transcript as a whole were extracted as separate text files (sub-corpora) and analyzed for word frequencies using a concordancing program (PhraseContext), which treats each sub-corpus as a database. This provides a count of how many times each speaker used each word. The two sub-corpora were then compared for differences in frequencies of use of the same words with what is called a log-likelihood calculation. This reveals which words were used with the greatest difference in frequency by the two speakers. Focussing on the differences in pronoun use, we find that Biden used the following personal and possessive pronoun forms more frequently than Palin did, in order of greatest difference first:

<table>
<thead>
<tr>
<th>Pronoun</th>
<th>Biden</th>
<th>Palin</th>
</tr>
</thead>
<tbody>
<tr>
<td>their</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>it</td>
<td>80</td>
<td>45</td>
</tr>
<tr>
<td>they</td>
<td>41</td>
<td>18</td>
</tr>
<tr>
<td>me</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>he</td>
<td>56</td>
<td>34</td>
</tr>
</tbody>
</table>

We see a predominance of third-person pronoun forms, which is less characteristic of spoken, interactional language. Surprising in this context is the frequent use of ‘me’ by Biden as compared to Palin. This seems to be partly due to rhetorical strategies he uses frequently, such as ‘let me’ (5 times, as in ‘Let me begin by…”) and story-telling (‘it didn’t take me long… to realize’), rather than, for example, to scenarios of self-objectification, about things that happened to me.
A comparison of Palin’s most frequently used pronouns which differed from those of Biden is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Biden</th>
<th>Palin</th>
</tr>
</thead>
<tbody>
<tr>
<td>our</td>
<td>21</td>
<td>71</td>
</tr>
<tr>
<td>we’re</td>
<td>13</td>
<td>32</td>
</tr>
<tr>
<td>you</td>
<td>60</td>
<td>95</td>
</tr>
<tr>
<td>we</td>
<td>97</td>
<td>138</td>
</tr>
</tbody>
</table>

The difference turns out to focus on first-person plural and second-person pronouns, conforming to the expectations of conversational language. As a side note, the use of ‘I’ was statistically close for the two speakers (Biden: 132, Palin: 118). A qualitative look at the more extensive use of ‘we’ by Palin shows the extent to which it played a role in trying to engage the audience with her. For example, consider the following excerpt:

Palin:
Let’s do what our parents,
told us,
before we probably even got that first credit card.
Don’t live outside of our means,
We need to make sure that as individuals,
we’re takin’ personal responsibility through all this.

Her story set-up involves what ‘our’ parents told ‘us’, and so the presumed quote from the parents should be “Don’t live outside of your means,” yet Palin reframes it in terms of the inclusive scenario, talking about ‘our means’. In another instance the focus shifts and involves a change in who is actually represented by ‘we’:

Palin:
Darn right it was the predator lenders,
who tried to talk Americans into thinking that um,
it was smart to buy a $300,000 house,
if we could only afford a $100,000 house.
There was a deception there,
and there was greed,
and there is corruption on Wall Street.
And we need to stop that,

First there is a shift from third person to first person (Americans > we), with the ‘we’ here representing Americans as consumer-citizens, particularly ones with a lower economic status. After the scenario is impersonalized in the following utterances (‘there was’), she returns to ‘we’ as agents who need to stop that corruption. But it is not those same lower-income Americans who are going to stop Wall Street on their own; there is a shift in ‘footing’ here (Goffman 1981), whereby the scope of the ‘we’ has changed to focus on the government, presumably in the future with McCain and Palin in charge. Though it remains to be investigated further, the extensive use of ‘we’ by Palin may not only serve to help frame her arguments in a more spoken-language-friendly way, but may also be part of a larger strategic use of the first-person plural pronoun (Bull & Fetzer 2006) to fudge reference when needed.
so as to sound like she is including the audience, even when in terms of logical reference she is not.

- colloquial language (vocabulary and pronunciation)

Since this category involves close analysis of the audio portion of the debate video, a selection of excerpts of the debate were used (henceforth, the “text sample”). The selection consisted of one third of the 90-minute debate, specifically the first ten minutes, the middle ten minutes (beginning 40 minutes into the debate), and the last ten minutes. This text sample thus includes the important opening and closing statements, as well as answers to question on different topics at different points in the debate. As a rule the partially articulated words were “corrected” in the transcription done by the US Commission on Presidential Debates to their full forms, and so were retranscribed for the text selection based on the audio from the video.

We will begin with colloquialisms in word choice, which are underlined below.

**Biden:**
colloquial vocabulary:
folks on Main Street
**Look folks,**
I said, "I'd feel like a jerk."
my dad used to have an expression. He'd say, "champ,

**Palin:**
colloquial vocabulary:
Darn right
darn right we need tax relief
pretty much
it's gonna be OK.
Joe Sixpack,
you walk the walk; you don't just talk the talk.

The colloquial expressions from this sample of the debate differ per speaker not only in terms of quantity but also qualitatively. Biden’s few colloquial words are all terms of address, either to the audience (folks) or in quotations referring to himself. They represent a limited range of linguistic usage. Most of Palin’s colloquial vocabulary consists of evaluative expressions. Such elements are more complex in their semantic role in that they represent the speaker’s attitude toward what (in this case) she is saying. Interestingly, all are phrases, which also marks them as more integral elements of linguistic usage, with the longest example being a variation on an idiomatic slogan (walk the walk ... talk the talk), notably pronounced with a marked regional accent from the northern midwestern US (the “flat a” in walk and talk). Such phrases are also a more integral part of one’s discourse in that they are less easily replaceable with equivalents, as opposed to single-word items.

Let us turn to the two speakers’ use of colloquial pronunciation in the text sample, marked by partial versus full articulation of words. This includes contractions in which two words are combined, but simple contractions with is (e.g., it’s versus it is) were not included here due to their common occurrence beyond just informal contexts in spoken American English.
Biden:
colloquial pronunciation:
*it's gonna* come as our-- our security services have said, as recently as *a coupla* weeks ago
Yeah, I can.

Palin:
colloquial pronunciation:
we're *gonna* fight for America.
We're *gonna* fight for the middle-class, the only thing that they're ever *gonna* agree on
Government’s *gonna hafta* learn to be more efficient
And I'll *betcha*
and get down to *gettin’* business done
we're *takin’* personal responsibility through all this.
that the economy is *hurtin’* like it is.
That’s *a lotta* middle income average American families,
*to learn* a heck of a lotta good lessons through this and say

Again we find Palin using more colloquial pronunciations, articulations that are more characteristic of conversational talk, both in terms of partial articulations (e.g., *gettin’*) and word blends (*betcha* instead of *bet you*). These also sometimes co-occurred with colloquial word choices (as in *a heck of a lotta*).

**- syntactic structures**

Based on analysis of the debate excerpt, Biden showed several instances where his argument structure conformed to the more complex referential structures of written discourse than to the more easily parsed norms of spoken language use. Consider the following example from Biden’s closing statement. (The passage is transcribed in intonation units, as explained in the following section.)

Time code on the YouTube video: 1:30:04

Biden:

We measure progress in America, based on whether or not, someone can pay their mortgage.
Whether or not they can send their kid to college. Whether or not, they're able to, when they send their child, like we have, abroad, or I'm about to, abroad, and John has as well, I might add, ... to fight.
That they are the best equipped, and they have everything they need.
This excerpt, which only takes about 20 seconds, involves a complex set of embedded clauses which, on first hearing, are dizzying, and which taken together do not make sense. The structure is as follows, with embedded clauses indented to show the layering.

Biden:

Whether or not, they’re able to,
   when they send their child, like we have,
   abroad, or I’m about to, abroad,
   and John has as well, I might add, ...
That they are the best equipped,

The utterance appears to be “whether or not they’re able to fight,” but this is because so many embedded clauses came before the main verb (fight) that Biden emphasizes the verb when he gets to it, which makes it sound like the end of the entire sentence/idea (when they send their child abroad to fight). In fact he the main verb (presumably something like ensure as in whether or not they’re able to ensure that they are the best equipped) gets left out.

Palin, by contrast, rarely uses such complex embedded constructions, but more often infers logical connections between ideas by uttering one short phrase after another, often with repetition of a verb and or other construction, e.g.,

But again, with some of these dictators, who hate America, and hate what we stand for, with our freedoms, our democracy, our tolerance, our respect for women’s rights, those who would try to destroy what we stand for,

This kind of repetition is more characteristic of spoken than written discourse.

In many utterances by Palin the subject is also clear because it is stated explicitly, even if it repeats words just uttered:

We'll know when we're finished in Iraq, when the Iraqi government can govern its people, and when the Iraqi security forces, can secure its people.
Such patterns are almost chant-like in their formulation, mechanical to the point of even producing grammatical infelicities (its people rather than the correct form their people, which would break the mold of the linguistic template she has created).

- **intonation patterns**

Intonation was annotated for selected sections of the text sample of the debate. The following are a representative example for each of Biden’s and Palin’s debate talk.

Intonation units were distinguished as per Chafe (1994) and Du Bois et al. (1993). The criteria involved work as a family resemblance category, such that units in which all the features co-occur are clearly discernable as discrete units, while for some speakers some features play a greater role than others. The five main features are pitch contour, change in loudness, acceleration followed by deceleration, change in duration (syllable-final lengthening), and alternation with pausing. As per the transcription practice in this notation, a period/full stop indicates an intonation unit ‘final’ contour (usually dropping to the bottom of the speaker’s pitch range), a comma marks a non-final intonation contour (one which indicates that the speaker will continue), a question mark shows high final rising intonation which is used as an appeal to the addressee, and an intonation unit which is truncated (cut short) is marked with two dashes (--). These marks therefore do not serve the function of traditional punctuation in written English.

Intonation and stress within each unit were transcribed using a variation on the Tone and Break Indices (ToBI) system (http://anita.simmons.edu/~tobi/). Syllables are marked below if there was prominent stress (greater volume), with greatest stress marked with an asterisk (*). Three basic tones were marked within a speaker’s voice range as exhibited in the debate: H for high, M for medium, and L for low (the binary high/low distinction of the ToBI system was adapted here into a three-way division for greater explanatory clarity). Tones which were glided to but not stressed (boundary tones) have been marked in parentheses (L). The time codes on the video clip (http://www.youtube.com/watch?v=89FbCPzAsRA) have been marked below for the beginning of each excerpt so that readers can listen to the accompanying audio.

The first transcript is of Biden at 1:23:05, in answer to the question, “Can you think of a single issue, policy issue, in which you were forced to change a long-held view in order to accommodate changed circumstances?” The following is how he begins setting up his answer.

Biden:

M* Yeah,

L* I can.

H* M*

uh when I got to the United States Senate,

M M L

uh and went on the Judiciary Committee,
as a young lawyer,
I was of the view,
and had been trained in the view,
that the only thing that mattered,
was whether or not,
a nominee,
appointed --
eh suggested by the president,
had a judicial temperament,

Here and for much of the debate, Biden uses a minimal extent of his intonational range, staying in the lower part of his voice range, and often the lowest part. This often results in a voice quality known as “creaky voice,” characteristic of low-energy speech, often occurring when a speaker is at the end of an intonation unit and nearly out of breath. For Biden, however, it is not limited to use in that circumstance but is a dominant pattern, resulting in a delivery which is rather monotone. In terms of stress, most of his intonation units have a two-part structure with two main stressed syllables. The difference between his more strongly stressed and regularly stressed syllables is minimal. Sometimes one of the two syllables per intonation unit is slightly higher in pitch than the other, but other times the two are both nearly the same pitch (M). Overall his use of intonation does not add much information to the speech, in terms of either marking discourse structure more distinctly or by showing the speaker’s attitude (affect) towards what he is uttering.

By contrast, let us compare part of Palin’s answer to a follow-up question to the one above, concerning how, as vice president, she might try to achieve greater bipartisan work between the two main political parties in Washington.

PALIN at 1:27:03
And, again,
voters on November fourth,

are gonna have that choice,

to either support a ticket,

that supports policies,

that create jobs.

You do that by,

lowering taxes on American workers and on our businesses,

and you build up infrastructure,

and you rein in government spending,

Overall, many syllables receive marked stress (increased volume, marked with the *) in Palin’s talk, and these highlight certain words against the background of the others. Typical of most of the debate, Palin also uses a wide range of pitch in her intonation, including great variation within a given intonation unit. Indeed, I used a L+ marking at one point to indicate a noticeable distinction made between L and M pitch accent in her vocal range. She also makes use of boundary tones, unstressed final pitches that are glided up to or down to. For example, several times she glides off of a higher stressed final syllable to a lower tone, which makes for a prosodically marked (and noticeable) ending of an intonation unit. She marks some lists (lowering taxes on American workers and on our businesses and and you build up infrastructure), with high rising boundary tones at the end, made more prominent by the length of the intonation units in words. Here this serves to set up the first two items on the list as a contrast with the third. Other patterns of sharply contrasting pitches within one intonation unit also mark parallel syntactic and semantic structures by using intonation contours which are parallel in structure (for the phrases that supports policies and that creates jobs), even if the second actually uses lower pitches overall than the first. Note that this provides cues to the listener as to how to interpret the words being spoken in terms of anticipating discourse structure. For example, the last two phrases mentioned here are part of a larger utterance, which could be rendered in a written sentence as “Voters on November fourth are going to have that choice to either support a ticket that supports policies that create jobs.” One would expect the next utterance to follow the either/or argument structure by introducing the “or” phrase, but Palin continues, “You do that by...” The “or” phrase does
come in, but only several intonation units later. However, the intonation contours with the parallel structures finish with a low (L) boundary tone, which is normally used to indicate the completion of an idea, making it sound like Palin has completed a point and so can go on to develop it (“You do that by...”). We can note that in contrast to Biden, many of Palin’s intonation units consist of three or four part (pitch) structures, with occasional extreme variation within one contour, as the last line of transcript above shows. In general her stress and intonation mark off not only the ideas that she is presenting but also the larger discourse structure, and sometimes the highlighting of the discourse structure can provide an audibly logical framing for certain phrases whose elements do not logically cohere.

- direction of eye gaze

The excerpt used above for the analysis of intonation is repeated here and extended, but is marked for eye gaze. Biden uses two gaze positions in this excerpt: looking down in front of him (marked with vvv) and looking up at the moderator of the debate (marked with ^^^).

Biden:

vvvv
Yeah,
vvvvv
I can.

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
uh when I got to the United States Senate,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
uh and went on the Judiciary Committee,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
as a young lawyer,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
I was of the view,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
and had been trained in the view,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
that the only thing that mattered,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
was whether or not,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
a nominee,

vvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvvv
appointed --
eh suggested by the president,

had a judicial temperament,

uh had not committed a crime of moral turpitude,

and was--

had been a good student.

And it didn't take me long --

It was hard to change,

but it didn't take me long,

but it took about five years,

for me to realize that,

the ideology of that judge makes a big difference.

That's why I led the fight against Judge Bork,

Biden's gaze rarely takes account of the viewer as addressee. His addressee, whom he looks at occasionally, is the debate moderator, who is off to his right side. In some respects, his gaze pattern is like that of conversation – looking away from the addressee (down) when formulating his ideas, and back at her when finishing expressing them. This occurs at the end of some intonation units, and more so towards the end of his turn than at the beginning. But the gaze is not at the television audience, and to complicate matters, his gaze at his addressee is off to the side. In alternation with his gaze downward to the space in front of him, the result is that his shifts in gaze can give an appearance of being “shifty-eyed,” and thus potentially suspicious. In addition, the shifting eye gaze from center (looking down) to right (looking up) goes with a slight head shift back and forth, from center to right and back, resulting in an ongoing horizontal head shaking pattern, similar to the one used by speakers of English when indicating ‘no’. Otherwise, his head movement is subtle, and supports the semantic valence of his assertions: generally head nods with positive comments, head shakes
with negative ones. His facial expression changes little, except for an occasional eyebrow raise for emphasis.

The excerpt for Palin transcribed above begins with her looking directly at the camera for the first several intonation units, so we will begin near the end of the excerpt and extend it, beginning at time code 1:27:15 on the video. Gaze at the camera is marked with 000, gaze to her right with <<< and gaze downward with ‘v’.

Palin:

And you build up infrastructure,

and you rein in government spending,

and you make our --

uh our nation energy independent.

Or you support a ticket,

that supports policies,

that will kill jobs.

by increasing taxes.

And that's what the track record shows,

is a desire to increase taxes,

increase spending,

a trillion-dollar spending proposal that's on the table,

That's gonna hurt,

our country,
Palin’s eye gaze engages the camera and thus the television viewer as the addressee. But it does more. Brief looks downward mostly occur at the beginnings or ends of intonation units. Since those intonation units are ones which consist of complete clauses, the glance punctuates an idea unit which consists of a proposition. The gaze pattern is therefore less like that of natural conversation, but serves to help parse her speech into intonation units/idea units. Furthermore, her prolonged gaze at the camera is consistent with her pattern of answering questions with the audience (you) as the subject. As she says in her closing statement of the debate, she likes speaking directly to the audience: “I like being able to answer these tough questions without the filter of the mainstream media” (regardless of the fact that the debate was broadcast on a major American television network).

Palin does look occasionally to her right side, and this is when she looks at Biden. The conspicuous move away from looking at the camera can serve to direct her audience’s attention at those moments to her opponent as well. It is noteworthy that Palin does this specifically when mentioning her opponents’ policies that are being presented with a negative evaluation. Overall her use of gaze therefore helps frame the structure of the ideas being presented and sometimes offers the added information of showing the source of proposals coming from the Democrats via Biden.

In addition to gaze, an additional layer of analysis could be performed on the candidates’ facial expressions. In brief it can be noted here that Palin made use of a variety of expressions at the camera, smiling much of the time, but also sometimes sneering when reporting proposals by the Democrats that she disagreed with; Biden maintained a consistent expression of seriousness most of the time. Finally one could even analyze the use of eye-blinking. Since Palin was looking at the camera most of the time, her blinking pattern was conspicuous, and often served to punctuate her speech with emphasis, whereas Biden’s blinking behavior was less discernable, particularly with his frequent looking downward.

Discussion:
We have seen different ways in which one politician in particular, namely Palin, made extensive use of the affordances of spoken communication in the setting of a televised debate. This concerns (at least) the levels of word choice (lexicon), pronunciation (phonetics), grammatical formulation (syntax), use of intonation (prosody), and co-verbal behaviors (such as eye gaze, head gesture, and potentially manual gesture as well). The more developed version of this paper will consider some of the reactions to this debate in the press just after it took place for comparison between the analysis here and the popular perception of the two candidates’ “performances.”

Taking advantage of these affordances in the context of self-presentation on television is part of a broader process. The increasing “mediatization” of politics (Mazzoleni and Schulz 1999) is now a familiar phenomenon, and one could argue that the use of spoken language
framing is a central (but not the only) process in this phenomenon. Mediatized spoken discourse does not simply adopt conversational norms, but adapts them to the context of the one-way media stream of television, radio, or web pages – although the latter involve separate consideration, due to the interactive possibilities of responding in online fora with text or video messages. The one-way format is what allows Palin to make a point that lacks logical coherence (but which follows a coherent intonation structure for an argument) and not immediately be questioned by her interlocutor as to what she meant. In the debate format, such critical response is usually delayed until the moderator takes her turn, by which point the follow-up question for clarification is diffused into a critique spread over the two debaters for not answering the question more directly.

Palin’s talk has been analyzed here as an example, but the phenomena discussed above are not unique to her. One of the contemporary masters of the process of mediatization of political discourse, and of oneself, is the Italian Prime Minister Silvio Berlusconi. Fedel (2003) discusses his use of such devices as:

- repetition with slight variations;
- the presentation of arguments through parataxis: simply juxtaposing ideas in short statements, rather than connecting them with coordinating conjunctions (and, or, but) or subordinating conjunctions (because, since, while); and
- the consequent reduction of reality to dichotomies, which allows for an X versus Y presentation (which can easily be framed in terms of us versus them).

It would be interesting to look more systematically at the televised speaking behaviors of politicians from different political parties in different countries to see if there is any pattern as to which ones adhere more closely to mediatized norms of conversational discourse patterns. If Palin and Berlusconi are any indication, I would hypothesize that this form of self-presentation is closely tied to populism, rather than to a particular left or right, or progressive or conservative, orientation. Through this kind of spoken language framing, the message can become more personalized toward the speaker him/herself, and as such take the argument away from ‘logical’ reasoning about issues to the level of personal and emotional persuasion. But how is spoken language framing effective in this regard?

**Implications:**

One process which this type of discourse framing can facilitate is ‘fictive interaction’ (Pascual 2008 and elsewhere). In the context of the type of political debate studied here, the speaker is not actually engaged in a dyadic encounter with an addressee like one would be in a conversation. Even a speaker who talks directly to the moderator, as Biden does, does not expect the turn-taking procedures of a conversation to be in play. For a speaker like Palin, who is addressing the camera, the one-sided nature of the talk is even more exaggerated. However, the devices of spoken language framing create a situation for the television viewer whereby it is as if the speaker were interacting with him or her. This is supported by the inclusive first- and second-person pronouns (we, our, you), the use of colloquial vocabulary and pronunciation reflecting an informal conversation, and the direct eye contact via the camera and television – whose indirectness, mediated by this technology, is put in the background in the context of this simulated direct encounter. (Note that Palin’s excuse to the moderator at one point when she was reminded that she had not answered a question was that she wanted to “talk straight to the American people.”) The use of repetition of phrases with slight variations is what often takes place in the back-and-forth exchange of conversations,
yet here it is the televised speaker him/herself who plays the role of providing the repetitions with variations, standing in for the contributions by the non-present addressee. The viewer is, or can be, drawn into a fictive personal encounter with the politician. In such a situation, the subjective experience of the encounter is heightened as objective distance is reduced. Sympathetic viewers can ignore the fact that an utterance by Palin began with “either…” and did not lead to an “or…” if they are riding the wave of her intonation pattern which makes for a convincing-sounding delivery (one can feel as if she has made a good point). With a speaker like Biden, who is mostly talking to a third person (the moderator) off to the side, there are fewer cues to engage viewers and take them out of the objective viewing stance. There is less support for fictive interaction and less grounds for understanding the speaker’s words on an emotional level, especially if s/he is using complex embedded clauses.

The notion of a simulated interaction with the speaker has an additional dimension on the cognitive level. A series of studies beginning in the mid-1990s has shown that there is a subset of neurons in the brains of humans and monkeys which responds when an individual performs certain actions (such as picking up a cup) as well as when the individual observes others performing the same movements (Rizzolatti et al. 1996). These nerves, now known as ‘mirror neurons’, provide a kind of simulation of the action which one is observing, even though s/he is not doing it. Though this simulation is below the level of conscious awareness, it provides a form of direct internal experience, and therefore a form of understanding of what performing that action would be like. Additional research has revealed ‘audio-visual mirror neurons’, ones which are activated not only when seeing someone perform an action but also when just hearing the sound of that action (such as ripping a sheet of paper). It has been suggested that mirror neurons help us understand others’ actions, emotional states, and even intentions (Rizzolatti et al. 2006). Given this background, we can see that the audio-visual cues of spoken language framing provide a richer basis for simulating a conversation with the speaker who uses them than with one who conforms more to the reading of a written text, for example. The natural use of emotive facial expressions, conversational intonation patterns, and other aspects can prime the viewer to be in the state of mind of a conversational encounter, possibly leading to greater affinity with the speaker. Like the cues of fictive interaction, this could also heighten the viewers’ affective response, making them more likely to respond to emotion-based logic.

In terms of research on spoken political discourse, the findings here indicate the need to take spoken language data on its own terms. Written transcripts are partial not only in the selectivity of what they represent of the actual language used, but they also increase the appearance of objectivity of the discourse and of how it may have been perceived and interpreted in real time.

Acknowledgement:
I am grateful to Gianluca Giansante for his discussion with me of Fedel (2003) and Mazzoleni & Schulz (1999).

References


