

The GoToGo construction: a descriptive study
Laura Staum
Stanford University

Introduction

This paper is an investigation of the construction that has been called the GoToGo construction, exemplified below:

1. *We're going to the American people and tell 'em the truth; that the unions are wanting to protect their Washington power.*
2. *Tomorrow you are going back to the store and apologize for stealing from him, and you're going to pay him for the things you tried to steal out of your allowance, do you understand?"* (Google search results)

This construction (hereafter GTG) is judged acceptable by some subset of American English speakers, but certainly not all or even a majority of them. Conducting a basic Google search for examples like this using frequent verbs returns hundreds of examples; due to the nature of the construction, these searches must be done with specific verbs, so it is impossible to know how many examples of this type exist, but it is certainly well enough attested for us to be sure that it is a part of the dialects of some native speakers of English. Anticipating results from the experimental section of this paper, grammaticality judgments indicate that approximately 20% of the surveyed population found sentences containing GTG within the realm of acceptability (see appendix for details on the sample and judgments).

GoToGo owes its name to Arnold Zwicky (Zwicky 2002). While its name is not as mnemonic as the names of some other constructions as employed in this paper, it does bear a relationship to some characteristics of the construction. First, the most commonly accepted instances of the construction involve the word *go* in both its motion and future

senses. Second, the name *GoToGo* evokes one of the possible sources of the construction, a type of sentence in which both future and motion *go* appear:

3. I'm going to go to the bank and deposit a check.

In this example, *go* appears twice; the first *go* is the future *go*, and the second is the motion *go*. In GTG, these two *gos* are telescoped into one instance of the verb, which conveys both meanings:

4. I'm going to the bank and deposit a check.

GTG seems to bear striking similarities to some other constructions in American English, all of which are consistently judged as grammatical by nearly all speakers of this variety:

5. I'm going home to get an umbrella.
6. I'm going home and getting an umbrella.
7. I'm going to get an umbrella.

These constructions are all prospective, in that they represent events that are understood to take place in the future, if they take place. Most of them also involve motion, with the notable exception of the third construction, which we will call *ProspectiveBeGoingTo* (hereafter *PBGT*).

The first two of the above constructions, which we will call *GoThereToV* (hereafter *GTTV*) and *GoThereAndV* (hereafter *GTAV*), respectively, can be interpreted as having approximately the same meaning as GTG, as below:

8. I'm going to the bank to deposit a check.
9. I'm going to the bank and depositing a check.
10. I'm going to the bank and deposit a check.

However, most syntactic analyses of GTTV and GTAV would be different, and as we will see in the next section, there are some semantic differences between them as well. This leaves us with several interesting questions to consider about the construction exemplified in the third sentence. How do people use it? What do we know about it? What are some possible analyses and how can we choose among them? What can we guess about its history/future?

In this paper I will describe some of the semantic properties of the construction, suggest a syntactic analysis, provide some data about its usage, and make some arguments about its possible sources and diachronic path.

Semantic Properties

Referring back to the three examples from the previous section (repeated below), our first task in an attempt to characterize GTG is to differentiate somehow among the three different constructions. Though I have stated that all three have approximately the same meaning, it is not true that they mean exactly the same thing or are appropriate for exactly the same situations.

11. I'm going to the bank to deposit a check.
12. I'm going to the bank and depositing a check.
13. I'm going to the bank and deposit a check.

Purposiveness

There are two semantic differences between 11 and 12. The first is that in 11 the purposive relationship between the first event (going to the bank) and the second event (depositing the check) is not optional – there is no way to interpret the sentence without it. In 12, there are two possible readings: one in which there is a purposive relationship between the events, and one where there is not. The default interpretation seems to be the

purposive one, if it is available. However, this relationship arises from a conversational implicature. If information is added that contradicts this implicature, it is canceled and the other interpretation is forced:

14. &¹I'm going to Wells Fargo to deposit a check at Bank of America.
15. I'm going to Wells Fargo and depositing a check at Bank of America.
16. & I'm going to Wells Fargo and deposit a check at Bank of America.

In this case, example 15 does not contain an implication that the first event should lead to the next or that the purpose of going to Wells Fargo is to deposit a check at Bank of America. However, since the implication is not optional in example 14, it is infelicitous with contradictory information (such as that Bank of America is not at Wells Fargo, and that going to Wells Fargo would not aid in the attempt to deposit a check at Bank of America). Preliminary judgments suggest that example 16 is also infelicitous, meaning that, like GGTV, the GTG construction is inherently purposive².

Presuppositions

These constructions also differ in their presuppositions about the beliefs of the speaker regarding event₂ (depositing a check):

17. I'm going to the bank to deposit a check, but I think it will be closed.
18. &I'm going to the bank and depositing a check, but I think it will be closed.
19. &I'm going to the bank and deposit a check, but I think it will be closed.

In these examples, the speaker thinks the bank will be closed and therefore that the check will not be deposited. This contention is contradictory in example 18, because GTAV

¹ I will use the ampersand sign (&) to indicate that an example is infelicitous; that is, it may be grammatical, but it does not make sense or does not convey the intended sense.

² The term "purposive" is being employed here as though it has an obvious, universally accepted meaning, which it does not. The meaning most applicable to the current circumstances has not yet become clear; this issue will be taken up again later.

implies that the speaker thinks the second event will take place (that the check will be deposited). However, example 17 is felicitous because GTTV sentences contain purposive phrases, which do not carry such an implication (there is no contradiction between the fact that the purpose of the trip was to deposit the checks and the belief that the speaker will not do so). Example 19 (GTG) seems to match with example 18 (GTAV) in this case, and is also infelicitous (according to some speakers – this has not been tested with a large sample). Example 20 demonstrates that PBGT shares this feature with GTG:

20. &I'm gonna³ deposit a check, but I think the bank will be closed and I won't be able to.

Below is a slightly formalized expression of the semantic properties of these constructions described above:

$P = X \text{ goes to the bank}$ $Q = X \text{ deposits a check}$ $W = \text{want}(x, P \rightarrow Q)$ ⁴

GoingThereToV

I'm going to the bank to deposit a check.

$P \wedge W$

"X goes to the bank and X wants the fact that X goes to the bank to lead to the eventuality that X deposits a check."

GoingThereAndVing

I'm going to the bank and depositing a check.

$P \wedge \text{then } Q$

"X goes to the bank and then X deposits a check."

³ "Gonna" is used here in place of "going to" to enforce the reading of PBGT rather than the possible reading of GTTV without a goal.

⁴ According to this representation, purposives are essentially resultatives with a modal operator.

GoToGo

I'm going to the bank and deposit a check.

$(P \wedge \text{then } Q) \wedge W$

“X goes to the bank and then X deposits a check and X wants the fact that X goes to the bank to lead to the eventuality that X deposits a check.”

ProspectiveBeGoingTo

I'm gonna deposit a check.

Q

“X deposits a check.”

From the above description of the semantic structures of these constructions, it seems that GTG is semantically a blend of GTTV and GTAV. That is, it contains both propositions P and Q, like GTAV, but it also contains W, like GTTV. What this means in practical terms is exactly what we have elucidated using examples earlier in this section: GTG requires that both verbs be in a realis mood⁵, like GTAV, but it also includes some notion of purposiveness (as expressed by W), like GTTV. The question that immediately presents itself in the face of these semantic similarities concerns the syntax of GTG: Is the syntax of GTG similarly parallel to these two constructions? The next section is devoted to considering the answer to this question.

Syntactic Analysis

The goal of our syntactic analysis is to determine a syntactic structure for GTG that would support the semantic structure noted above. We will start by considering several possible analyses of GTG that are suggested by its parallels to other constructions.

⁵ I take this to mean that the speaker must believe that the proposition is true, as modified by the tense/aspect of the clause.

Conjunction

One possible analysis of GTG is as a conjunction, much like GTAV, that simply has a morphologically exceptional second conjunct. This analysis has the benefit that it gives us P \wedge Q for free; that is, the nature of conjunction gives us the interpretation that both verbs in a GTG construction would have realis mood⁶. However, as we will see in the discussion of a complementation analysis, it has several syntactic disadvantages; in addition, it has the semantic disadvantage that it fails to account for the fact that GTAV and GTG have different relationships to W.

Complementation

Another way to analyze the GTG examples is to consider the second VP embedded in the first. This would distinguish it from GTTV examples, which are standardly interpreted as having purposive phrases adjoined to the main VP. If we call the second VP an adjunct when it is a purposive phrase (as in GTTV), but a complement in PBGT and GTG, then in our analysis *and* in GTG functions like *to* in PBGT, in that it takes verbal complements, making it a complementizer.

This would make GTG syntactically similar to PBGT and very different from GTTV; additionally, it would make GTG the only way to syntactically encode realis purpose in English (filling in a paradigm of sorts).

In order to determine whether this analysis is a plausible one for the constructions under discussion, it is necessary to consider just what differentiates a complement from an adjunct and a conjunct, and also to determine what the consequences would be of adopting this analysis.

⁶ This mechanism involves the tense of the first verb in the case of GTG; if the first verb is in the present progressive (which we interpret as future), and the conjunction tells us that the second event takes place after the first event, then we can infer that the second event takes place in the future as well (and that it does, in fact, take place).

Complement vs. Adjunct/Conjunct

The Cambridge Grammar of the English Grammar (Huddleston and Pullum 2002) proposes several tests to differentiate between complements and adjuncts. Below I've selected the three that can (somewhat) fruitfully be applied to the constructions in question. Though there are not typically tests for distinguishing between complements and conjuncts (this not normally being a confusing distinction in English), we will see that at least one of them may help us in making this distinction as well.

Licensing

Complements require a licensing verb; adjuncts can appear with any verb (they are less restricted). Here, we apply this test to purposive clauses, GTG V2 clauses, and PBGT.

Purposive clauses can attach to any verb (unrestricted).

- I'm cleaning my room to please my mother.
- I'll bring a book to read.

GTG V2 Clauses, on the other hand, can only attach to *going* for some speakers, motion verbs for other speakers, etc.

- I'm going to school and get an education.
- ?*I'm cleaning my room and please my mother.
- ?*I'm eating a hamburger and get fat.

PBGT only exists with the verb *going*.

- I'm going to get an education.

Obligatoriness

Complements are sometimes obligatory; adjuncts are always optional. Purposive clauses are always optional, making them look like adjuncts.

- I'm going to the bank (to deposit a check).

This test does not really apply to GTG V2 clauses; if we try to apply it, we get the same sentence as above:

- I'm going to the bank (and deposit a check).

But we could also achieve this result from GTAV:

- I'm going to the bank (and depositing a check).

This may come down to something akin to polysemy – *going* has multiple lexical entries, some of which subcategorize for a verbal complement, others of which don't.

The infinitive phrase in a PBGT construction is required; when it is absent, the sentence gets interpreted as ellipsis.

- I'm going to...

Anaphora

Complements must all be included in the referent of an anaphor such as “do so,” but adjuncts don't have to be. Purposive clauses are not included, again supporting an analysis of these as adjuncts:

- I'm going to the bank to deposit a check; would you like to do so to open an account?

This test doesn't work very well for GTG examples, because “do so” is part of a reasonably formal register, which makes it inappropriate for use with a GTG construction (which is quite informal), so all sentences sound “funny” for this reason, making it difficult to determine whether they are “grammatical” or even whether they make sense:

- ?&I'm going to the bank and deposit a check; would you like to do so and open an account?

For prospective be going to, the only possible reading of the sentence below is that the addressee is being invited to deposit a check for the purpose of opening an account, so the infinitival phrase is clearly a complement:

- &I'm gonna deposit a check; would you like to do so to open an account?

Overall, as Table 1 shows, it is difficult to find tests that apply to GTG, but the evidence seems to mildly support an analysis in which the second VP in GTG examples are considered complements of the first VP.

Table 1. Complement or Adjunct?

	GTTV (purposive)	GTG (purp/prosp)	PBGT (prospective)
Licensing	Adjunct	Complement	Complement
Obligatoriness	Adjunct	?	Complement
Anaphora	Adjunct	?	Complement

Though none of these tests were intended to be relevant to conjunction, upon reflection, the licensing test provides some evidence against the conjunction analysis as well. Like adjuncts, conjuncts do not need to be licensed; this is a peculiarity of complements alone. In the GTAV examples, for instance, the first verb can, in fact, be anything:

- I'm going to the bank and getting some groceries.
- I'm eating some grapes and drinking some juice.

They must be pragmatically sensible in order for the pragmatics to contribute any purposive semantics, but they are not ungrammatical without the purposiveness. Any two verbs can be conjoined this way. If GTG were correctly analyzed as a conjunction, we would expect it also to be free of restrictions on the conjuncts. Yet as the test above demonstrated, only going (for some speakers) or only motion verbs (for other speakers) license the second VP in GTG examples. There may be speakers who have no restrictions on what the first verb in a GTG construction may be; this would certainly be an interesting finding, but it does not change the fact that for many speakers, a conjunction analysis is impossible because of these restrictions.

Analyzing the GTG examples as instances of complementation is motivated by some other syntactic facts. One is that, aside from the present progressive forms that GTG examples are most commonly found in, this construction can appear in some other environments where only –ing forms are allowed, as in the past progressive:

21. I was going to the bank to deposit a check when...
22. I was going to the bank and depositing a check when...
23. I was going to the bank and deposit a check when...

and as gerundives:

24. I hate going to the bank to deposit a check.
25. I hate going to the bank and depositing a check.
26. I hate going to the bank and deposit a check.

Though not all speakers find examples like C acceptable, they were easy to find on Google:

27. *I think in times like this, going to school and see your most hateful teacher is kinda better.* www.dragon-girl.net/ladyluck/archives/2003_04.html collected 2004 from Google

28. *Makes me feel like going home and get a trishaw ride.* www.penangturfclub.com/penang/esplan/esplan.html collected 2004 from Google

Because GTG forms contain verbs that are not –ing forms, the two verbs cannot individually satisfy the morphological requirements for these positions (as the verbs in 25 do), so they must satisfy them together as one main verb with a verbal complement (as the verbs in 24 do).

Formalization

To make the complementation analysis clearer, it may be useful to look at a more concrete representation of these constructions:

GoToGo

=[gO_{pres.part.(motion/prospective)} + PP + VP_{inf/conj}]
=[going + [PP] + [Compl_{inf/conj} + VP_{base}]]
=[going + [PP] + [and + [V_{base} + [complements of V]]]]
[going [to the bank] [and [deposit [a check]]]]

Here, the verb *go* in its present participle form, with both motion and prospective meaning, combines with a prepositional phrase and a VP to form the GTG construction. The construction stipulates that the VP be of the form *inf/conj*, which is intended here to mean that it must be infinitival, but must be introduced by a conjunction. This VP consists of a complementizer of the form *inf/conj* plus a base form VP; the only available complementizer of this form is *and*. This property is intended to distinguish the VP in GTG from the VP in PBGT of the form *inf* below.

ProspectiveBeGoingTo

=[be + [gO_{pres.part.(prospective)} + VP_{inf}]]
=[be + [going + [Compl_{inf} + [VP_{base}]]]]
=[be + [going + [to + [V_{base} + [complements of V]]]]]
[be [going [to [deposit [a check]]]]]

Here, the VP of the form *inf* is composed of a complementizer of the form *inf* plus a base form VP. The default complementizer of this type is *to*, so in this construction the VP is introduced by *to*.

In the last two constructions, GTTV and GTAV, the verb *go* can appear in any form, not just the present participle. Also, only the motion meaning is required; though tense/aspect features can give it a prospective/future type of meaning, it does not carry this meaning itself. In GTTV, we see the same version of VP_{inf} that we saw above.

GoThereToV

= [go_(motion) + PP] + [VP_{inf}]
 = [go + [PP]] + [Compl_{inf} + [VP_{base}]]
 = [go + [PP]] + [to + [V_{base} + [complements of V]]]
 [go [to the bank]] [to [deposit [a check]]]

GoThereAndV

= [go_(motion) + PP] + Conj + [VP]
 = [go + [PP]] + and + [V + [complements of V]]
 [go [to the bank]] and [deposit [a check]]

As the representations above are intended to show, GTG shares all aspects of its structure with PBGT, with the exception of an added PP as a goal. GTTV, on the other hand, has a different structure in which the second VP is not embedded in the first. GTAV, as a further contrast, has a completely different structure in which *and* functions as a conjunction.

And as a complementizer

If, as the above discussion suggests, the best way to analyze GTG constructions is as a complex VP containing a VP complement, then we must consider the place of *and* in this structure. If our parallel to PBGT in structure holds, then *and* would be analyzed as having the function of an infinitival complementizer, just as *to* does in PBGT.

(I have chosen to phrase this analysis in reference to the function of *and* rather than to the category of *and* because I am specifically trying to comment on syntactic properties of expressions here, rather than syntactic properties of words, which is what category labels refer to.)

If we analyze GTG as involving complementation rather than conjunction or adjunction, then we are left to find some other way to explain why the second verb has realis mood. My proposal is that while *and* is functioning as a complementizer in these cases, it retains its status as a conjunction in category. It is well known that when words have syntactic functions that do not match their syntactic category, they can exhibit some behaviors characteristic of both. Some work regarding this idea can be found in the recent volume *Mismatch: Form-Function Incongruity and the Architecture of Grammar* (Francis and Michaelis 2002) and in dissertations by both Malouf and Wescoat (Malouf 1998, Wescoat 2002). In the case of GTG, the *and* is behaving in its capacity as a complementizer by introducing a verbal complement, and it is behaving in its capacity as a conjunction by providing the information that the second verb follows the first verb temporally (giving both “conjuncts” equal status with respect to mood).

This behavior is suspiciously reminiscent of a change in progress; that is, it would not be unexpected for us to find *and* behaving as a complementizer with no trace of its status as a conjunction in a later stage of English. Noonan (1985) notes that such a grammaticalization path (from conjunction to complementizer) is cross-linguistically attested. A discussion of how this might have happened/be happening follows in the section on the diachrony of GTG. Before we can approach this question, however, I would like to provide some data about the construction and how it can be used that may shed further light on it.

Usage Data

In order to investigate speakers' assessments of sentences involving GTG, I conducted an experiment to collect acceptability judgments.⁷ I chose to use data based on judgments rather than usage because the construction is not all that common, and a researcher could grow old waiting for people to produce it. I used corpus methods to find examples on which to base some of the ideas in this paper; however, there simply weren't enough examples to be able to do any kind of quantitative work. These methods could provide me with evidence that speakers can produce certain types of constructions, but because of the small sample, there was no way for me to know if the gaps I observed were accidental or if they signified something. I decided to conduct a study of judgments in order to fill in some of this missing information and to get a better picture of what distinctions subjects make among different types of GTG examples.

Subjects participated in an online experiment where the task was to judge the acceptability of 72 sentences (of which 30 were non-experimental 'filler' sentences) on a scale of their own determination (see appendix for details on experimental materials and procedures). Approximately 120 subjects participated in the experiment. Subjects read a set of instructions and then took part in a practice session to familiarize themselves with the interface and the task before rating the sentences (which appeared in a partially randomized order).

The sentences subjects judged were set up in minimal pairs or near minimal pairs; there were several variables involved in the design of the experiment, but only the following variables are being considered here:

Variable	Variants
Verb	going, coming, running
Goal	PP, home, fishing, no goal

⁷ I avoid the term "grammaticality judgments" here because I feel that the term "acceptability" better captures the assessment I asked the subjects for – see appendix for the instructions subjects received.

The experimental sentences about which the data for this paper were gathered all contained a GTG construction that involved one of the above verbs (*going*, *coming*, or *running*). In the position directly following the verb, there was a locative goal (the word *home* or a prepositional phrase), the adverbial⁸ *fishing*, or nothing. One of the goals of the study was to determine whether all GTG sentences are equally acceptable to a GTG speaker or if the speakers distinguished between the sentences depending on which of the above verb/goal variants they contained.

Because I was primarily interested in the distinctions drawn by subjects who found the general paradigm of GTG sentences acceptable, the data presented here is based on a subset of the total subjects that was identified as being sympathetic to the experimental sentences (using a clustering technique to assign subjects to this group). This subset consists of slightly less than half of all subjects (47); a large number of subjects gave all GTG sentences such a low rating (comparable to ratings they gave to sentences that are not a part of any native American English speaker's repertoire) that I consider them to be categorical non-speakers of the GTG dialect, if there is such a thing. For this reason, I did not want to consider their intuitions, because they are not likely to tell us much about how GTG is actually used.

Results

From the raw judgments that people reported for each sentence in the experiment, I calculated normalized judgments that are comparable both within a speaker and between speakers (see appendix for details about the normalization procedure and the treatment of the data). I was most interested in determining which sentences were consistently judged to be more acceptable than others, and seeing if there were different patterns of acceptance exhibited by different subjects.

⁸ The analysis of the –ing form in a sentence such as “I’m going fishing” as an adverbial is taken from Silva 1975. It is notable that all of the other elements described as “goals” in the current study (such as locatives) are also adverbial in nature.

The main set of results that I will discuss here concerned both of these issues. In order to convert the ratings into a binary variable (for ease of analysis), I established a cutoff point for the normalized judgments of 0.25, such that any subject whose average judgment was less than 0.25 for a given set of sentences was counted as considering these sentences unacceptable, and any subject whose average judgment was greater than 0.25 was counted as considering these sentences acceptable. It is not important exactly what level of acceptability this cutoff point represents; it is only important that we know that it represents the same level of acceptability for all subjects (which is ensured by the normalization procedure). If the pattern established by this method would also apply at different levels of acceptability, that is an additional finding, but it does not have any bearing on how robust this pattern is.

The sets of sentences that I considered were those GTG sentences that shared a certain value for one of the variables (for example, all the sentences containing the verb *going* or all the sentences containing *fishing* as the goal). Because the data was structured as sets of near-minimal pairs, these sets of sentences were extremely similar for each variable (see appendix for list of sentences). I used the subjects' average judgments over these sets as their judgment for these variables. Once I had divided the subjects into "acceptable" and "unacceptable" groups for each set of sentences, I calculated the conditional probability that any subject who found one value for a variable acceptable would find another one acceptable⁹.

Verbs

As you can see from the table below, there are certain pairs of variables that show very high conditional probabilities. In some cases we can see that a high probability

⁹ The conditional probability represents the chance that if one condition holds another condition will hold. This was determined by calculating the proportion of subjects for whom one of the sets was acceptable for whom another of the sets was acceptable. For example, 14 subjects found *coming* acceptable and of those subjects, 13 of them also found *going* acceptable, so the conditional probability for this pair is 13/14.

represents a correlation that is not directional; for example, in the case of *coming* vs. *running*, both directions have a reasonably high conditional probability. So if a subject considered *coming* acceptable, he or she has a 10/14 (71%) chance of finding *running* acceptable, and the chances of someone who found *running* acceptable finding *coming* acceptable are almost exactly the same (77%). This means that the groups who find them both acceptable overlap to a great extent, which is an interesting finding itself, but it does not mean that either of them is more broadly acceptable than the other, because the total numbers of people who find each acceptable are almost exactly the same (13 and 14).

However, in the case of *going*, the number of people who find it acceptable is much larger (22). A simple two-tailed t-test tells us that this represents a significantly different proportion of the 47 subjects than 13 or 14 (which are not significantly different from each other); the probability that the coming results were only randomly different from the going results was $p=0.0135$, and the probability of the opposite, that the going results were only randomly different from the coming results, was $p=0.0042$.

The fact that more subjects accept *going* than the other two verbs leaves us with two possibilities for its relationship to the other verbs: it can either be more broadly acceptable than *coming* and *running*, or it can simply be acceptable to a larger but independent group of people. As the table below shows, the probability that someone will find *going* acceptable if he or she finds *coming* or *running* acceptable is very high (93% and 85%, respectively). This implies that it is the first case that is demonstrated by the data: *going* is more broadly acceptable than *coming* or *running*. Another way of putting this is that those subjects who judged *coming* or *running* to be acceptable are basically a subset of those who judged *going* to be acceptable. An interesting supporting fact is that of the 22 subjects who rated sentences with *going* as acceptable, only two of them rated *coming* and *running*, on average, higher than *going*. Based on these observations, it seems fairly safe to say that GTG sentences involving the verb *going* are more acceptable to GTG speakers than those involving the verbs *coming* and *running*, and they are also acceptable to a wider spectrum of speakers.

Table 2. Conditional Probabilities - Verbs

If you find X acceptable	then there is a Y/Z chance that	you also find W acceptable
going	13/22 (59%)	coming
going	11/22 (50%)	running
coming	13/14 (93%)	going
coming	10/14 (71%)	running
running	11/13 (85%)	going
running	10/13 (77%)	coming

I performed several chi-square tests to determine the significance of these findings; below are the results of the chi-square for going vs. coming.

Table 3.

going by coming				
going	Coming			
	Bad	Good	Total	
Bad Frequency	24	1	25	
Cell Chi-Square	2.3677	5.5811		
Good Frequency	9	13	22	
Cell Chi-Square	2.6906	6.3422		
Total	33	14	47	
Statistic		DF	Value	Prob
Chi-Square		1	16.9816	<.0001

The bottom right cell represents those subjects who rated both values of the variable as good; this is the figure we have been using for comparison above. The p-value that is the result of the chi-square obviously represents the probability that we would get the entire distribution as a result if the null hypothesis were true and there were no association between how subjects rate sentences with going as the verb and sentences with coming as the verb. This has clearly been rejected, as the p-value is the lowest value that the statistical package will report¹⁰. The number in bold in one of the four cells is the highest cell chi-square value, which means that the cell containing it contributed the most to the

¹⁰ The statistical analysis was done using SAS. See appendix for details.

significance of the distribution. As we will see, this cell turns out to be the good/good cell for all of the verb combinations, which tells us that the good/good cell has a greater impact on the association (which is highly significant) than any other cell. Below are the results for going vs. running, which are very similar to the going vs. coming results above, though slightly less significant:

Table 4.

going by running			
going	running		Total
	Bad	Good	
Bad Frequency	23	2	25
Cell Chi-Square	1.3357	3.4934	
Good Frequency	11	11	22
Cell Chi-Square	1.5178	3.9697	
Total	34	13	47
Statistic	DF	Value	Prob
Chi-Square	1	10.3166	0.0013

The results of the run for coming vs. running (Table 6) are even more significant than the first run; however, as discussed above, this tells us only that the same subjects find coming and running acceptable. It does not indicate any sort of implicational relationship between these two verbs. It is interesting, though, that there is no such relationship – coming appears in many constructions with going that do not allow other motion verbs (such as the GoAnd construction discussed in the diachronic section), but GTG does not seem to privilege the verb coming over running at all. Because the expected count in one of the cells of Table 6 is less than 5, the p-value that emerges from the Fisher’s Exact Test, below, is more accurate than the result of the chi-square. The p-value is in bold.

Table 5.

Fisher's Exact Test	
Two-sided Pr <= P	4.021E-05

Table 6.

coming by running			
Coming	running		Total
	Bad	Good	
Bad Frequency	30	3	33
Cell Chi-Square	1.5729	4.1137	
Good Frequency	4	10	14
Cell Chi-Square	3.7075	9.6965	
Total	34	13	47
Statistic	DF	Value	Prob
Chi-Square	1	19.0906	<.0001

Goals

Among the different goals, we can see a similar pattern of conditional probability. Here, however, it is the *no goal* variant that has the most predictive power. This seems to be the most restricted case. Unlike the verb situation, in which there is one general case and two restricted ones, in the case of the goals, there is one restricted case (the *no goal* variant, which has nine subjects who judge it to be acceptable) and three more general cases (all of which have 20-24 subjects who judge them to be acceptable). Here, all subjects who find the *no goal* variant acceptable (that is, a sentence like “I’m going and get a hamburger”) are virtually guaranteed to find the other three variants acceptable.

Table 7. Conditional Probabilities - Goals

If you find X acceptable	then there is a Y/Z chance that	you also find W acceptable
PP	18/25 (72%)	home
PP	17/25 (68%)	fishing
PP	9/25 (36%)	No goal
home	18/21 (86%)	PP
home	15/21 (71%)	fishing
home	9/21 (43%)	No goal
fishing	17/20 (85%)	PP
fishing	15/20 (75%)	home
fishing	8/20 (40%)	No goal
No goal	9/9 (100%)	PP
No goal	9/9 (100%)	home
No goal	8/9 (89%)	fishing

Even within the small population of nine subjects who find the restricted case of no goal acceptable, there are distinctions to be made. Four of the subjects found all three types of goals more acceptable than no goal. One subject made no distinction among the four possibilities (the three goals and no goal). The remaining four subjects found at least one of the goals less acceptable than no goal (though often not by very much, and these differences are very unlikely to be significant). What we do know, though, about these four subjects is that, like the one subject who did not distinguish at all, they certainly do not categorically find sentences with goals significantly more acceptable than sentences with no goal. So overall, we have three types of speakers: those who require a goal (all but the nine discussed above), those who do not distinguish on the basis of presence of a goal (the last five mentioned above), and those who find sentences with goals better but still accept sentences without goals (the first four discussed above).

I also performed a chi-square to test the significance of the goal results. Below we see the results for *home* vs no goal:

Table 8.

home by no goal			
home	no goal		Total
	Bad	Good	
Bad Frequency	26	0	26
Cell Chi-Square	1.1792	4.9787	
Good Frequency	12	9	21
Cell Chi-Square	1.4599	6.1641	
Total	38	9	47
Statistic	DF	Value	Prob
Chi-Square	1	13.7820	0.0002
Fisher's Exact Test			
Two-sided Pr <= P		2.157E-04	

This data shows a slightly different pattern from what we saw in the verb section. The highest cell chi-square appears in the cell representing subjects who found sentences with *home* good and sentences with no goal good. This association contributed the most to the highly significant result; what we can gather from this is that it is especially striking how many subjects were in this cell (the expected value was much lower than the observed value). That is, no goal examples are more predictive of PP examples than the other way around (so PP is more widely acceptable). Below are the results of PP vs. no goal:

Table 9.

Table of PP by no goal			
PP	no goal		
	Bad	Good	Total
Bad Frequency	22	0	22
Cell Chi-Square	0.9978	4.2128	
Good Frequency	16	9	25
Cell Chi-Square	0.878	3.7072	
Total	38	9	47
Statistic	DF	Value	Prob
Chi-Square	1	9.7958	0.0017
Fisher's Exact Test			
Two-sided Pr <= P		0.0019	

Here, though the numbers are very small, we still achieve strong significance because the association is so strong. Again, the top right cell has the highest cell chi-square. As before, we interpret this to mean that no goal examples are predictive of PP examples (so PP examples are the general case). The results for no goal vs fishing are similar:

Table 10.

Table of no goal by fishing			
no goal	fishing		
	Bad	Good	Total
Bad Frequency	26	12	38
Cell Chi-Square	0.7966	1.0755	
Good Frequency	1	8	9
Cell Chi-Square	3.3636	4.5409	
Total	27	20	47
Statistic	DF	Value	Prob
Chi-Square	1	9.7767	0.0018
Fisher's Exact Test			
Two-sided Pr <= P	0.0026		

Analysis

Based on the above data and some examples from a variety of corpora, it is possible to put together a coherent story about the current state of GTG usage and its possible past and future. This is not an attempt to explain how or why these changes might take place or have taken place; that will be left for the following section. In this section, I am most interested in providing a description of the changing grammatical landscape for speakers. The path suggested here is supported by both the existence/non-existence of certain types of examples in corpora from different time periods and by frequency of acceptability in the current study. This “typology” is meant to suggest that for each type, there may be speakers who accept or reject it, and for each speaker, all examples that they accept will be differentiated by type from all examples they reject. However, there are further distinctions to be made than those I am making here, that will be discussed in the section on further work.

Type 1

Verb: *going* only

Goal: explicit locative

Examples:

29. *I'm going back and tell Terry and Gottlieb they can go to the devil...* (1925 S. Lewis, Arrowsmith (Grossett & Dunlap) xxvii.300, from David Denison)
30. *I am going to the sanatorium and get my wife and daughter and quit this place forever.* (www.artandfantasy.com/cursefleshstorypage10.html - collected 2004 from Google)

Type 2

Verb: *going* or *coming*

Goal: explicit locative

Examples:

31. *I'm coming over there and drag you out myself.* (1934 *It Happened One Night* [movie] from DD)
32. *Someday, I'm coming over and see how the hell you make this stuff.* (www.deviantart.com/view/6585458/ - collected 2004 from Google)

Type 3

Verb: any motion verb

Goal: less clearly locative, still required

Example:

33. *I'm taking him to the Sheriff and make sure he's destroyed.* (1939 N. Longley, F. Ryerson, E. A. Woolf, *Wizard of Oz* [film dialogue] from DD)

Type 4

Verb: not necessarily motion verb

Goal: not necessarily locative

Example:

34. *I'll be turning the key and see if it works.* (1997 phone conversation from DD)

Type 5

Verb: *going*

Goal: optional

Example:

35. *I worked on a little while longer and then I decided well, I'm going and get my Master's Degree.*

(www.llf.lib.ms.us/Winnebago/LLF/Oral%20H...ries/HARRIS.htm collected 2004 from Google)

In a sense, this typology represents a description of GTG as it can currently be used; that is, there are speakers for whom some of these types are acceptable and not others, and so for each subset of speakers, some portion of this typology is descriptively accurate. There are, of course, types of possible GTG constructions not represented in this typology. I have included here only those constructions whose acceptability I was able to test to some degree in my experiment.

The Diachrony of GoToGo

Aside from the question of how the different types of GTG constructions may have developed out of one another, there is the obvious question of how the construction emerged as part of some speakers' grammars in the first place. Both of these questions are the subject of this section.

Reanalysis

GTG is not the only construction in English that allows *and* in some suspiciously complementizer-like positions. It is superficially very similar to what I will call the GoAnd construction (hereafter GA); they both have optional locative goals (for at least some speakers):

36a. Why don't you go and deposit a check? (GA)

36b. I'm going and deposit a check. (GTG)

37a. Why don't you go to the bank and deposit a check? (GA)

37b. I'm going to the bank and deposit a check. (GTG)

They appear to be very structurally similar, which brings us to an interesting question: is *and* in GA functioning as a complementizer? At first glance, the answer seems to be no:

38. I'll go and open the door.

39. I went and opened the door.

This appears to be a simple conjunction, with no complementation in sight. The conjunction analysis is also capable of explaining why the second verb in GA also carries the implication that the second event is expected to be accomplished:

40. You can go to open the door, but it's stuck.

41. &You can go and open the door, but it's stuck.

This implies that GA is also realis, which licenses this implicature; however, it doesn't tell us whether it is a conjoined main verb or it is a verbal complement. This is, in a sense, a syntactic ambiguity, which could lead to some speakers having different interpretations of this than other speakers.

This is exactly the situation we are looking for to explain why there are some speakers who find GTG examples grammatical and some who don't. If some speakers analyze GA examples as simple conjunctions, they do not have *and* functioning as a complementizer available to them as a resource with which to create or analyze GTG examples. However, if some speakers choose to analyze GA examples as main verbs with verbal complements (presumably by analogy to GoTo examples), then they are equally capable of creating GTG examples by the same analogy to PBGT examples.

The process of reanalysis described above has been documented in a variety of situations, including the development of the French inflectional future and the English modal auxiliaries (Hopper and Traugott 1993). Hopper and Traugott describe the process thus: “A hearer has heard the ‘output’ ... but assigns it to a different structure after matching it with possible... structures” (1993 41). The result of reanalysis, they explain, is “the development of new out of old structures,” and this is exactly what we see in the case of GTG (1993 56). Hopper and Traugott point out that as long as ambiguity remains, there is no way to determine that reanalysis has taken place. It is only when unambiguous versions of the new structure emerge that it is clear that a reanalysis has taken place. Here, GTG seems to be the unambiguous new structure that clues us in to the reanalysis that took place in the ambiguous environment of the GA construction.

One problem with using this ambiguity as the source of the GTG construction is that if it is truly based on a reanalysis of GA examples, we would expect the original GTG examples to be those without explicit locative goals, because the GA construction is much more commonly used without such goals; however, as we have seen, GTG examples seem to be most acceptable to GTG speakers with the locative goals, and these examples are much more frequent than the examples without goals. However, it is important that we recognize that GA can still be a source of GTG in the sense described above without GTG actually being modeled after GA. If, as suggested above, GTG examples are an analogy to PBGT and are simply allowed by the newly available function of *and* as a complementizer, then they ought to look more like PBGT examples on the surface, while simply using a new resource made available by a reanalysis of GA examples.

Hopper analyzes GA as an example of hendiadys; that is, he recognizes that these are not, in fact, straightforward examples of coordination, but decides to analyze the first ‘conjunct’ as an auxiliary, rather than analyzing them as main verbs with verbal complements taking *and* as a complementizer. This certainly seems to be a reasonable analysis, and it does not seem totally incompatible with the story given above. If the first

verbs of hendiadys constructions are auxiliaries, then it is left to a theory of auxiliaries to determine whether *and* is a complementizer in this structure, and there is no obvious reason why it could not be one.

Telescoping

As discussed in the introduction to this paper, one possible source for the GTG construction is the existence of similar sentences in which each meaning of *go* is represented by a separate instance of the verb:

42. I'm going to go to the bank and deposit a check.

Viewing this type of sentence as a source for GTG is not incompatible with the above story regarding reanalysis. The telescoping that would have to take place to turn the above sentence into a GTG example might be encouraged by the existence, for some people, of *and* functioning as a complementizer. Before this is available, the most likely interpretation of example X would be the following:

[I [am [going [to [go [to the bank]] and [deposit [a check]]]]]]]

Once *and* has become available functioning as a complementizer, example X could be interpreted as follows:

[I [am [going [to [go [to the bank] [and [deposit [a check]]]]]]]]]

If this interpretation, in which the *deposit* VP is subordinate to the *go to the bank* VP is used by speakers, it becomes possible to eliminate the first instance of *go* without violating the grammatical principle that two conjuncts in English must be of the same morphological form.

Seed Hypothesis

If we approach the GTG construction as one that represents a change in progress, these results support the “seed hypothesis” (Zwicky 2002), in which constructions are spread through certain high frequency examples, and these examples provide the model for generalizations. These high frequency exemplars tend to maintain their own high frequencies, so that even though the construction may have become acceptable in a wider domain, the original seeds will often still be the most frequent examples. In this case, it seems that the seed form of the GTG construction involved the verb *going*, and for some speakers, these examples are still the only ones that are acceptable. For other speakers, however, the examples with other motion verbs are also acceptable, though for many of these speakers they are not given judgments as high as those given to the *going* examples. There are, however, some subjects for whom any motion verb seems to be equally acceptable.

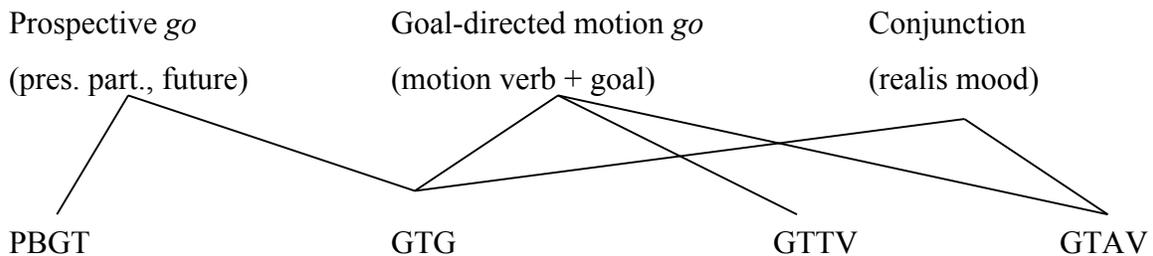
Cognitive Grammar

Another possible model for understanding the development of GTG is something similar to the Cognitive Grammar model used by Tuggy (1996) to derive what he calls the 2-B construction (also referred to as *Isis*, *is-is*, *double-be*, etc). This model suggests that ‘sanction’ for new constructions can be given by the existence of a family of similar constructions, some of whose characteristics can be combined in a new way to create a construction that was previously ungrammatical. As we have discussed at length in this paper, there are several such constructions available to sanction the development of GTG; all of its characteristics could be traced to one or more of these already grammatical constructions. While the Cognitive Grammar model, as described by Tuggy, does not provide a coherent explanation for the process by which this ‘sanction’ and development of new constructions occurs, it does indicate that there are other documented examples of such construction development, and that however we choose to model it, the introduction and spread of a previously unknown construction that is related to other established constructions is one of the ways in which language change takes place.

Multiple Inheritance

Ginzburg and Sag 2000 provide a way of formalizing the way that constructions can inherit attributes from multiple sources. In Figure 1 we see a possible way to describe the relationships among GTG and related constructions. The diagram shows that GTG inherits some of its characteristics from several parent constructions; as Ginzburg and Sag point out, “by organizing phrases into a multiple inheritance hierarchy, one can posit higher-level types and formalize the relevant cross-cutting generalizations” (2000, 7). Below, we see that GTG shares future meaning and the requirement that it be in the present participle with PBGT, it shares motion meaning and the necessity of an adverbial goal with GTTV and GTAV, and it also shares its realis mood with GTAV.

Figure 1. Multiple Inheritance Diagram for GTG and related constructions



Suggestions for Further Work

In addition to the types discussed above, there are uses of GTG that I have not considered that deserve further investigation. Some of these are attested, but I have no information on whether they are generally acceptable, and some of them are unattested, and I would like to determine whether they are really ungrammatical for all speakers.

Attested: Gerunds

Some speakers/writers of GTG produce examples in which the GTG construction appears as a gerund. These examples are not extremely common, but they seem to be perfectly acceptable to at least some people. I know of at least a few people for whom other GTG constructions are acceptable who do not find these acceptable. I think including examples like these in an experiment similar to the one with which the data for this paper was gathered might allow me to determine whether these examples fit into the typology of GTG constructions, and if so, where.

43. *I think in times like this, going to school and see your most hateful teacher is kinda better.* (www.dragon-girl.net/ladyluck/archives/2003_04.html collected 2004 from Google)

44. *Makes me feel like going home and get a trishaw ride.*
(www.penangturfclub.com/penang/esplan/esplan.html collected 2004 from Google)

Unattested: Non-motion examples

While I did gather some data (much of it unanalyzed; see below) involving non-motion verbs, I have not seen any examples involving going that do not involve motion. That is, there are no GTG examples in which going has only prospective (future) meaning, though this meaning is obviously attested by PBGT. I find these examples totally ungrammatical myself, but I suspect that they may be grammatical to some speakers, or, if they are not now, they may become grammatical to some speakers in the future. I would be interested in gathering some data on how people assess these examples as well.

45. *I'm going and give you some advice right now.

46. *I'm going and die right here.

Additional Data

In addition to the data discussed in the paper so far, the experiment yielded a lot of data that has not yet been analyzed. Some of it involves further variations on the variables discussed, and some of it involves other questions that have not yet been addressed.

- Non-motion verbs
- Non-purposive examples
- Examples involving adverbs
- Examples of TryAnd

In the future, I plan to analyze some of this data to determine whether there is any relationship between judgments of TryAnd sentences and GTG sentences, whether GTG sentences can involve verbs that do not expressly involve motion, etc. Another thing I would like to pursue is the question of whether those speakers whose judgments were not considered in this analysis exhibit similar patterns to those discussed here, or whether “non-speakers” of GTG reveal different patterns of judgments than “speakers” of GTG.

Additional Analysis

In addition to further exploring data that has not been considered in this paper, I would like to consider different ways of analyzing the data presented here. Specifically, I am interested in taking advantage of the gradient nature of the judgments collected by using these judgments themselves instead of converting them to categorical judgments and analyzing them using ANOVA. This would have a few major benefits. First, it would allow me to see if the patterns that appear using the categorical version of the data are upheld by the gradient data, and to what extent. Second, it would allow me to investigate the possibility of interactions between the variables. This benefit is especially important, because though using chi-square for the analysis of this data might indicate that I do not expect the variables to interact, on reflection I think there is a reasonably good chance that the verbs and the goals will show some evidence of interaction.

Because of the large amount of data collected overall in the course of this study, I expect that many further interesting findings may emerge upon further analysis. Though the social information collected was limited, information on age, gender, and occupation/education was collected for many subjects, and it is also possible that this would be an interesting avenue for research.

References

- Denison, David. 2002. Personal Communication with Arnold Zwicky.
- Francis, Elaine J. and Laura Michaelis. 2002. *Mismatch: Form-Function Incongruity and the Architecture of Grammar*. CSLI.
- Ginzburg, Jonathan and Ivan Sag. 2000. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. CSLI.
- Hopper, Paul. 2002. Hendiadys and Auxiliation in English. *Complex Sentences in Grammar and Discourse: Essays in honor of Sandra A. Thompson*. Joan Bybee and Michael Noonan (eds.). John Benjamins.
- Hopper, Paul and Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge UP.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge UP.
- Malouf, Robert. 1998. *Mixed Categories in the Hierarchical Lexicon*. Stanford University.
- Noonan, Michael. 1985. Complementation. In *Language typology and syntactic description. Vol II – Complex Constructions*. Timothy Shopen (ed). Cambridge UP.
- Pullum, Geoffrey K. 1990. Constraints on intransitive quasi-serial verb constructions in modern colloquial English. *OSUWPL* 39.
- Pullum, Geoffrey K. and Arnold Zwicky. 1991. Condition Duplication, Paradigm Homonymy, and Transconstructional Constraints. *BLS* 17, 252-66.
- . 1999. Gerund participles and head-complement inflection conditions.
- Silva, Clare M. 1975. Adverbial –ing. *Linguistic Inquiry*, Vol 6.2. MITPress.
- Tabor, Whitney. 1994. Syntactic Innovation: A Connectionist Model. Dissertation, Stanford University.
- Tuggy, David. 1996. The thing is is that people talk that way. The question is is Why? *Cognitive linguistics in the redwoods: The expansion of a new paradigm in linguistics*. Eugene H. Casad (ed.), 713-52. Berlin: Mouton de Gruyter.
- Zwicky, Arnold. 1990. What are we talking about when we talk about serial verbs? *OSUWPL No. 39*. OSU Department of Linguistics.
- . 2002. *Seeds of Variation and Change*. NWAV 31 - handout.

Acknowledgements

I would like to thank everybody, both in the department and outside it, who has helped me with this project; most specifically, thanks go to the following individuals:

John Beavers and Liz Coppock for very fruitful discussions

Florian Jaeger, for invaluable assistance with the design and implementation of the experiment

Roger Staum, for serving as a consultant on matters relating to data and statistics

Arnold Zwicky, for being helpful, encouraging, smart, and fun

Appendix

1) Subject Pool

Subjects were recruited for participation in the experiment via an email request for participants that was sent to friends, family and acquaintances of members of the Stanford Linguistics Department. Of those who received this email, all of those people who wished to participate were accepted as subjects; I specified clearly that only native speakers of American English were acceptable subjects, but I did not attempt to verify any of the subjects' language backgrounds in any way. Because I used this method of subject selection, the results of this experiment cannot hope to provide information about any well-defined population, and cannot be generalized over any such population. The goal of this experiment was not to determine the characteristics of any population regarding their use of GTG; rather, it was an attempt to access the intuitions of a fairly large pool of native speakers of American English in the hopes that some of these speakers might have interesting intuitions on the subject of GTG examples. Thus the composition of the subject pool was of little concern so long as it contained at least some speakers who accept GTG sentences enough to discern among different types of them.

2) Interface

All subjects participated in the experiment using a web interface designed and maintained by Florian Jaeger. This interface can be viewed at the following URL:

symsys.stanford.edu:8081/experiment/web_exp3/me.going.index.html

3) Instructions

The instructions each subject encountered before taking part in the experiment appear below:

Personal Details

As part of this experiment, we have to collect a small amount of personal information, which you should enter in the Personal Details window once the experiment is started. *This information will be treated as confidential, and will not be made available to a*

third party. None of the responses collected in this experiment will be associated with your name in any way. If you have any questions about this practice, please contact the experimenter.

Filling in the Personal Details questionnaire at the start of the experiment is *not* required, but appreciated. In order to continue with the experiment the software requires that you fill in *something* in each field (note: the 'AGE' field will only accept number inputs and the 'EMAIL' field requires something of the form 'x@y.z'). This information will not be given to anyone and most of it will not even be used for research purposes. It solely helps to distinguish the different files from different participants in this study and will be deleted immediately after the data has been gathered. If you prefer you can fill in nonsense information. Only the age, gender, and region fields may be used in the analysis.

The following fields will appear:

- your name and email address;
- your age and sex;
- whether you are right or left handed (based on the hand you prefer to use for writing);
- the academic subject you study or have studied (or your current occupation in case you haven't attended university);
- region: please specify the place (city, region/state/province, country) where you lived the longest between the ages of 5 and 15.

Instructions

Part 1: Judging Line Length

Before doing the main part of the experiment, you will do a short task involving judging line length. A series of lines of different length will be presented on the screen. Your task is to estimate how long they seem by assigning numbers to them. You are supposed to make your estimates relative to the first line you will see, your *reference* line. Give it any number that seems appropriate to you, bearing in mind that some of the lines will be longer than the reference and some will be shorter.

After you have judged the reference line, assign a number to each following line so that it represents how long the line is in proportion to the reference. The longer it is compared to the reference, the larger the number you will use; the shorter it is compared to the reference, the smaller the number you will use. So if you feel that a line is twice as long as the reference, give it a number twice the reference number; if it's a third as long, provide a number a third as big as the reference.

So, if the reference is this line, you might give it the number 10:

If you have to judge this line, you might assign it 17:

And this one might be 2.5:

There is no limit to the range of numbers you may use. You may use whole numbers or decimals. If you assigned the reference line the number 1, you might want to call the last one 0.25. Just try to make each number match the length of the line as you see it.

Parts 2 and 3: Judging Sentences

In Part 1 of the experiment you used numbers to estimate the length of lines on the screen. In Parts 2 and 3 you will use numbers to judge the acceptability of some English sentences in the same way.

You will see a series of sentences presented one at a time on the screen. Each sentence is different. Some will seem perfectly OK to you, but others will not. Your task is to judge how good or bad each sentence is by assigning a number to it.

As with the lines in Part 1, you will first see a *reference* sentence, and you can use any number that seems appropriate to you for this reference. For each sentence after the reference, you will assign a number to show how good or bad that sentence is in proportion to the reference sentence.

For example, if the reference sentence was:

(1) The dog the bone ate.

you would probably give it a rather low number. (You are free to decide what 'low' or 'high' means in this context.) If the next example:

(2) The dog devoured yesterday the bone.

seemed 10 times better than the reference, you'd give it a number 10 times the number you gave to the reference. If it seemed half as good as the reference, you'd give it a number half the number you gave to the reference.

You can use any range of positive numbers that you like, including decimal numbers. *There is no upper or lower limit to the numbers you can use, except that you cannot use zero or negative numbers. Try to use a wide range of numbers and to distinguish as many degrees of acceptability as possible.*

There are no 'correct' answers, so whatever seems right to you is a valid response. We are interested in your first impressions, so please don't take too much time to think about any one sentence: try to make up your mind quickly, spending less than 10 seconds on each sentence.

Specific information on the experiment

All examples in the experiment are supposed to represent spoken American English. Please keep this in mind as you are judging the sentences (and please imagine the sentences spoken with the most natural intonation possible). These sentences do not have to be acceptable in a written text for you to give them a high rating. The highest ratings should be assigned to the sentences you could definitely imagine yourself uttering, in any possible context. The lowest ratings should go to the sentences you could not imagine any native speaker uttering in any context. Ratings in between these should be assigned to those sentences that you might say but you are not sure, sentences you would probably not say but someone else would say, sentences you are not sure someone else could say, etc. Please try to distinguish as many levels of acceptability as possible.

Any two sentences do not have to be judged equally acceptable by you even if you think they might have had exactly the same words in the same order. Do not try to remember previous sentences for the sake of comparison. Just use your intuitions in each individual case.

Finally, note that some sentences - even though they may be 'grammatical' according to some descriptive grammar - will just sound worse to you than others. In that case, follow your intuition. If the only way to understand a sentence is nonsensical, then you *are expected* to assign a relatively low number to this sentence (that is, to judge it to be a 'bad' or 'not so good' sentence).

Procedure

First please fill in the Personal Details questionnaire as described above, and then press the Start button.

The experiment will consist of the following 3 parts:

- Training session: judging 6 lines
- Practice session: judging 6 sentences
- Experiment session: judging 72 sentences

In each part you will see the reference item in the experiment window. Please enter your reference number and then press the Continue button. The test items will appear one after the other in the experiment window. Please type your judgment in the box below each item.

The experiment will take 15 to 25 minutes. After the experiment is completed you will receive an email confirmation of your participation if you entered your email address in the personal information questionnaire.

Please keep in mind:

- Use any number you like for the reference sentence.
- Judge each sentence in proportion to the reference.
- Use any positive numbers which you think are appropriate.
- Use high numbers for 'good' sentences, low numbers for 'bad' sentences and intermediate numbers for sentences which are intermediate in acceptability.
- Try to use a wide range of numbers and to distinguish as many degrees of acceptability as possible.
- Try to make up your mind quickly, and base your judgments on a careful first impression.

Please be patient after clicking the button. It may take a few seconds to load the experiment.

4) Sentences

Below are the sentences that subjects were presented with to judge. The reference sentence they saw with every experimental item and against which they were instructed to judge the acceptability of other sentences was the following:

I'm going home and gotten scolded by my mom.

The experimental sentences are presented below, including those sentences on which the data for this study are based and those sentences whose data were not analyzed:

I'm going to the beach and get a tan.
I'm going home and get an umbrella.
I'm going and get a hamburger.
I'm coming to the beach and get a tan.
I'm coming home and get an umbrella.
I'm coming and get a hamburger.
I'm running to the pool and go swimming.
I'm running home and get an umbrella.
I'm running and get a hamburger.
I'm going to the beach and getting a tan.
I'm going home to get an umbrella.
I'm going to the beach and go swimming.
I'm coming over there and see you.
I'm going home and go swimming.

I'm going and find out what happened.
I'm coming home and go swimming.
I'm going fishing and catch ten salmon.

I'm turning the key and see if it works.
I'm taking him to the store and buy him some clothes.
I'm going to school and run into my ex-girlfriend.
I'm going to work and get yelled at by my boss.
With my luck, I'm getting pulled over and lose my license right away.
I'm going skiing and have an accident.
I hate going to the dentist and get my teeth cleaned.
I'm looking forward to going home and take a nap.

I'm going to school and get my degree.
I'm graduating and never go to school again.
I'm going to college and get a good job.

Try and don't do anything stupid.
Try and not do anything stupid.
Try not to be late.
Try and be quiet or you'll wake the baby.
Try hard and not do anything stupid.
Try hard to be on time.
Try hard and be quiet or you'll wake the baby.
Try hard and don't do anything stupid.

I'll try and get there by seven.
I'm trying and get there by seven.
I'll plan and get there by seven.
I'll plan to get there by seven.

Below are the “filler” sentences that were presented intermingled with the experimental sentences:

He tries and sings every day.
I'm coming over sit in the kitchen.
I don't know who he thinks that eats sushi.
I wonder which dog that was eating my steak.
I'm going to the movies with my sister.
I'm eating nachos in front of the television.
He's going to the party with his girlfriend.
She loves Frank Sinatra.
John is so tall to date.
I don't regard him tall.
It is himself that John said would go to the store.
She doesn't regard me very smart.

It didn't pan out to be true.
 I'm not going to play myself short tonight.
 How big of a dog is it?
 The war looks bad on the United States.
 This'll be an interesting one to see how it fits.
 What she is is smart.
 The problem is is she's always late.
 The thing is is my dad.
 My dog is is a Pomeranian.
 My problem was with libraries was like you always have to be quiet.
 The point I was trying to make was was the violence.
 That's what I hate, is that she's always late.
 He calls a spade a fucking spade, is what he calls it.
 This is where the first question becomes more interesting, is who speaks these dialects?
 I haven't answered your second question, is what is the point of all this?
 I wish I had my umbrella, is what I wish I had.
 Here's something you should consider, is whether you really like him.
 My mom's excited to hear our concert, which, by the way, we should start rehearsal.
 Let's drive to the grocery store, which, by the way, we need gas.
 It's important to trust each other, which I never really trusted Sarah completely.

5) Normalization procedure

Each subject's judgments of the experimental sentences were normalized in comparison to his or her own judgments of four sentences designated as "bad" and four sentences designated as "good" that were used as filler sentences during the experiment (below):

Table X. Sentences used for normalization.

BAD	GOOD
He tries and sings every day.	I'm going to the movies with my sister.
I'm coming over sit in the kitchen.	I'm eating nachos in front of the television.
I don't know who he thinks that eats sushi.	He's going to the party with his girlfriend.
I wonder which dog that was eating my steak.	She loves Frank Sinatra.

These sentences were chosen on the basis that they represent sentences that native speakers would not consistently produce (the bad sentences) and sentences that native speakers would consistently produce (the good sentences). I did not select word salad or completely nonsensical items for the “bad” sentences because I wanted to compare the subjects’ judgments of the experimental sentences with sentences that were comparably sensible, so that if the experimental sentences were judged to be better than these sentences, it would presumably be due to their form and not their content. However, I wanted these sentences to be ungrammatical to the extent that it would be difficult to imagine anyone saying them consistently, so that if the experimental sentences were judged to be worse than these sentences, there would be no question that such a judgment meant they were ungrammatical.

The normalization procedure took an average of the four “bad” sentences and an average of the four “good” sentences and used them in lieu of minima and maxima in a reasonably standard normalization procedure, so that the difference between each judgment and the “minimum” (ie average bad sentence) was divided by the range (average good – average bad) to give a proportion. A sentence that was judged to be exactly as bad as the average “bad” sentences would therefore receive a score of 0, and a sentence that was judged to be exactly as good as the average “good” sentences would receive a score of 1. A sentence that was judged to be in between these two would receive a score in between 0 and 1, and a sentence that was worse than the “bad” or better than the “good” sentences would receive a score outside this range. This allows us to compare judgments not only within one speaker but between speakers because the judgments are situated between a common level of “good” and “bad” on an absolute scale. While there is no way to know how acceptable or unacceptable the “good” and “bad” sentences are for any subject, we do know how acceptable or unacceptable the experimental sentences were to the subjects relative to a constant level – the sentences in Table X.

- 6) Clustering Technique for selecting subjects that were ‘sympathetic’ to GTG

In order to select the subjects whose data I used in the analysis, I used a cluster analysis technique (called FASTCLUS) to find subjects who were similar to each other based on their distance from a cluster mean. The following is a description of the technique used from the statistical package documentation:

The purpose of cluster analysis is to place objects into groups or clusters suggested by the data, not defined a priori, such that objects in a given cluster tend to be similar to each other in some sense, and objects in different clusters tend to be dissimilar. You can also use cluster analysis for summarizing data rather than for finding "natural" or "real" clusters; this use of clustering is sometimes called *dissection* (Everitt 1980).

The FASTCLUS procedure combines an effective method for finding initial clusters with a standard iterative algorithm for minimizing the sum of squared distances from the cluster means. The result is an efficient procedure for disjoint clustering of large data sets. PROC FASTCLUS was directly inspired by Hartigan's (1975) *leader algorithm* and MacQueen's (1967) *k-means algorithm*. PROC FASTCLUS uses a method that Anderberg (1973) calls *nearest centroid sorting*. A set of points called *cluster seeds* is selected as a first guess of the means of the clusters. Each observation is assigned to the nearest seed to form temporary clusters. The seeds are then replaced by the means of the temporary clusters, and the process is repeated until no further changes occur in the clusters. (from SAS documentation)

I used this method of choosing the subjects to include in my analysis to avoid biasing the results by selecting any specific level of acceptability to use as a cutoff point.

7) Statistical Analysis

The statistical analysis of the results of the experimental portion of the paper was done using SAS (Statistical Analysis System). I used the chi-square test to test the hypothesis of independence (using categorical data, rather than the gradient data collected), Fisher's Exact Test, which is a correction of the chi-square for small numbers, and the t-test to test the significance of differences among the numerical results.

The results of the t-tests that were not fully discussed in the body of the paper appear below. The first set of probabilities reflect the likelihood that, if the results for *going* represent the chances that a subject will like a GTG sentence and the verb does not change them, the results for *coming* would obtain (and vice versa). The next set of results is the same, but with the pair going/running, and the last set of results is the same for the pair coming/running (unsurprisingly yielding non-significance).

going/coming	Left Tail= 0.0131902806	Right Tail= 0.0003452174	Prob= 0.013535498
	First= 0.995546666	Second= 0.9997684594	Prob= 0.0042217934
going/running	Left Tail= 0.0057615945	Right Tail= 0.0001079629	Prob= 0.0058695574
	First= 0.9984398694	Second= 0.9999395619	Prob= 0.0014996925
running/coming	First= 0.6939530478	Second= 0.9999999997	Prob= 0.3060469519
	Left Tail= 0.445150127	Right Tail= 3.39742E-10	Prob= 0.4451501273

Below are the results of the chi-squares not reported in the paper:

Table of PP by fishing			
PP	fishing		
	Bad	Good	Total
Bad Frequency	19	3	22
Cell Chi-Square	3.2023	4.3231	
Good Frequency	8	17	25
Cell Chi-Square	2.818	3.8043	
Total	27	20	47
Statistic	DF	Value	Prob
Chi-Square	1	14.1476	0.0002

Table of home by fishing			
home	fishing		
	Bad	Good	Total
Bad Frequency	21	5	26
Cell Chi-Square	2.4618	3.3234	
Good Frequency	6	15	21
Cell Chi-Square	3.048	4.1147	
Total	27	20	47
Statistic	DF	Value	Prob
Chi-Square	1	12.9480	0.0003

Table of PP by home			
PP	home		
	Bad	Good	Total
Bad Frequency	19	3	22
Cell Chi-Square	3.8328	4.7454	
Good Frequency	7	18	25
Cell Chi-Square	3.3729	4.1759	
Total	26	21	47
Statistic	DF	Value	Prob
Chi-Square	1	16.1270	<.0001