Acquisition of the constraints on *wanna* contraction by advanced second language learners: Universal Grammar and imperfect knowledge

Robert Bley-Vroman
University of Hawaiʻi

Soo-Ok Kweon
Konkuk University, Seoul

Abstract

Contraction of *want to* to *wanna* is subject to constraints which have been related to the operation of Universal Grammar. Contraction appears to be blocked when the trace of an extracted *wh*-word intervenes. Evidence for knowledge of these constraints by young English-speaking children in as been taken to show the operation of Universal Grammar in early child language acquisition. The present study investigates the knowledge these constraints in adults, both English native speakers and advanced Korean learners of English. The results of three experiments, using elicited production, oral repair, and grammaticality judgements, confirmed native speaker knowledge of the constraints. A second process of phonological elision may also operate to produce *wanna*. Learners also showed some differentiation of contexts, but much less clearly than native speakers. We speculate that non-natives may be using rules of complement selection, rather than the constraints of UG, to control contraction.

Introduction: *wanna* contraction and language learnability

In English, *want to* can be contracted to *wanna*, but not invariably. As first observed by Lakoff (1970) in examples such as (1), in which the object of the infinitival complement of *want* has been extracted by *wh*-movement, contraction is possible, but not in (2), in which the subject of the infinitival complement is extracted from the position between *want* and *to*. We shall call examples like (1) "subject extraction questions" (SEQ) and examples like (2) "object extraction questions" (OEQ).

(1)   Object Extraction Question (OEQ)
        a. Who do you want to kiss *t*?
        b. Who do you wanna kiss?

(2)   Subject Extraction Question (SEQ)
        a. Who do you want *t* to kiss Bill?
        b. *Who do you wanna kiss Bill?

There are several accounts of this asymmetry. The most widely accepted explanations agree in attributing the blocking effect in SEQ to an intervening trace, indicated by *t* in the examples above (or by some equivalent "global" device that makes reference to the "underlying position' of an extracted *wh*-word). The dominant formulation is that of the trace theory of movement rules (Chomsky, 1976; Chomsky, 1977; Lightfoot, 1976; Lightfoot, 1977; Chomsky & Lasnik, 1978; Chomsky & Lasnik, 1978; Rotenberg, 1978; Chomsky, 1980; Fiengo, 1980; Jaeggli, 1980). In trace theory, the presence of a Case-marked trace blocks contraction. Only (not Case-marked) PRO intervenes between *want* and *to* in (1), so contraction is possible. But, in (2) the trace of *who* means that *want* and *to* are not adjacent in the relevant sense, so that contraction is blocked.

This account presumes that the source of the constraint is, ultimately, Universal Grammar as it interacts with the structure of English. Thus, evidence for the constraint in learners has been taken as evidence for the operation of Universal Grammar. For example, Crain and Thornton (1998) demonstrated that even very young children show a very clear asymmetry: the children they studied commonly contracted in object-extraction questions (over 80% of the time), but they rarely contracted in subject-extraction questions. Crain and Thornton argue that this result shows that Universal Grammar is active in child language acquisition.

Clearly, the question of whether the constraint also operates in adult language acquisition should have consequences for the ongoing study of the role of Universal Grammar in SLA, and—more generally—for research into those areas of language knowledge which are not readily derived from the input to which learners exposed. (See Eubank (1991) for a good overview of the issues.) If adult learners are constrained by the relevant UG principles, permitting examples like (1) but not those like (2), then, one could conclude that the learners' competence must make use of abstract representations involving empty categories (or some functional analogue to them).

Now, empty categories are by their nature not audible or otherwise directly observable in the language data to which learners are exposed; still less are distinctions among empty categories directly observable in the input. And, the evidence available to the learner that some types of empty categories block contraction while others do not is not robust (see footnote 4). Though *wanna* itself occurs frequently enough, the relevant contrasts between possible and impossible contraction are probably not common, and—more telling—the contrasts would seem to need the support of negative evidence of the sort that learners surely are not reliably exposed to, evidence that contraction is not possible in certain specific cases. These factors are the basis for a standard poverty-of-the-stimulus argument. The conclusion is clear: the relevant abstract properties and principles cannot be learned: they must be available to the learner in advance of and independent of the input, presumably via Universal Grammar.

If, on the other hand, adult foreign-language learners' acquisition is chiefly input-driven, relying on observable co-occurrence patterns in the language data which the learner encounters, then, given the abstractness and complexity of conditions and the lack of definitive evidence, the distinctions, if made at all, should be made imperfectly, evincing considerable uncertainty and variability.

There are many assertions about the grammaticality of *wanna*-contraction in native-speaker English in the published linguistics literature, and we do not question their accuracy. Published empirical studies regarding actual performance on *wanna*-contraction are scarce, however. This is true even with respect to the performance of adult native speakers. In the case of foreign language learning, there is virtually nothing. The Crain and Thornton study of children is suggestive. One might reasonably speculate that if the children make the distinction so clearly, using Crain and Thornton's methods, then surely native-speaking adults would do so as well.

In one study of adults that we are aware of (Karins & Nagy, 1993), native speakers preferred a subject interpretation of extracted *who* in the sentence *Which one would you want to help?* at a rate of about 80%. However, if *wanna* was presented instead of *want to*, the preference for subject interpretation was greatly reduced (to a bit over 30%)—a reduction consistent with the theory. However, the fact that about a third of the subjects gave an interpretation which conflicts with the constraint suggests at least that in a

comprehension experiment, what Karins and Nagy call "pragmatic skewing" can obscure the operation of syntactic constraints.[1]

The purpose of the present research is to investigate knowledge of the constraints on *wanna* contraction in adults learning English as a second language. The learner group consisted of adult native speakers of Korean learning English at an advanced level. Each of our three experiments uses a different method: elicited production, imitation (used as a technique to elicit repairs), and grammaticality judgement. We also investigated native-speaker performance on *wanna* contraction—both as a comparison, or "control" group and to contribute to filling in the gaps in our knowledge of the phenomenon in native speakers. We report here the results of three experiments investigating *wanna*-contraction. We describe the design of each experiment in turn, followed a general presentation of common elements of analytical logic. Then, we present our finding and discuss their implications for the theory of second language acquisition.[2]

Experiment 1: Elicited production

The elicited production experiment was closely modeled on that of Crain and Thornton (1998), where *wanna* was elicited from children in the context of a "game." As pointed out above, Crain and Thornton were quite successful in drawing out the structures under investigation. We too found that the method is very effective: in pilot work, we saw that our adaptation of the technique was virtually guaranteed to result in the production of the relevant structures.

In the present study, a doll named "Emily" was introduced to the participant. To provide a lead-in context, we recorded a short contextual narrative by "Robert", introduced as a mutual friend of Emily and the participant. Given this context, Robert then prompts the participant to ask Emily a question. The prompt uses a phrase like "Ask Emily who." Examples follow. The key parts of the interchange are in boldface.

Example protocol for eliciting object extraction questions (OEQ)

Robert [on tape]: We are cooking in the kitchen now. I am making *pulkoki*, my sister is making pizza, and my brother is making *capchae*.
Robert: **I bet Emily wants to help someone. Can you ask Emily who?**
Subject: **Who do you want to/wanna help?**
Emily [as spoken by the experimenter]: You!

Example protocol for eliciting subject extraction questions (SEQ)

Robert [on tape]: I am thinking about cooking *pulkoki* for Emily's birthday and my sister is, too. I think I am a better cook than my sister.
Robert: **I think Emily wants one of us to cook. Can you ask Emily who?**
Subject: **Who do you want to/*wanna cook?**
Emily [as spoken by the experimenter]: You!

We aimed to elicit sentences of about the same length in both the OEQ and in the SEQ conditions. In OEQ sentences, there is no overt direct object after the verb (because of *wh*-extraction): *Who do you want to help?* In SEQ sentences, our use of intransitive or optionally transitive verbs (e.g. *Who do you want to cook?*) ensured that there was no necessity for an overt NP object in the elicited sentence. The experimenter's prompt always consists of two sentences: a statement with *someone* or *something* and a truncated embedded *wh*-question. Note that neither Robert's context nor the experimenter's prompt contained any cases (or any potential cases) of contracted *wanna*. That is, the participant does not hear *wanna* at all. Any case of contraction must originate in the participant's own linguistic system.

The elicited production experiment consisted of 12 items, 6 for OEQ and 6 for SEQ. Items were randomized. There were two practice items before the main test. Responses were tape-recorded.

Experiment 2: Oral repair

The oral repair test is one of the general family of tests called "elicited imitation" in language acquisition research (see Bley-Vroman & Chaudron (1994) for a review). What makes it specifically oral repair, rather than elicited imitation more generally, is that the participants are told they may modify the sentences if they sound unnatural. The idea is that if *wanna* is

presented in an example in which contraction would be impossible for the participant, there is a likelihood that it will be replaced by *want to* rather than simply be repeated.

There were 10 tape-recorded *wh*-extraction questions, all containing *wanna*: 5 examples of OEQ and 5 examples of SEQ. All were of approximately the same length (13 words ± 1 word). Instructions were: "Please listen and repeat the following sentences. You can modify the sentences if they seem unnatural." Items were randomized. Responses were tape-recorded.

Experiment 3: Grammaticality judgements

Participants were given 10 sentences: 5 OEQ and 5 SEQ. All test sentences contained *wanna*. Examples are *What do you think they wanna see in Hawaiʻi?* (OEQ) and *Who do you think the boys wanna make pizza?* (SEQ).[3] Following the practice of other researchers (e.g. Hawkins & Chan (1997) and Munnich, Flynn, & Martohardjono (1994), each test sentence was presented simultaneously in written form and in (tape-recorded) spoken form. Participants saw each sentence on a separate piece of paper and at the same time listened to the sentence played on a tape recorder. Test sentences were presented with a five-second gap between them. Subjects recorded their evaluations of the "possibility" by making one of four choices: Absolutely possible, Probably possible, Probably impossible, and Absolutely impossible. Following the methodology of Bley-Vroman, Felix, & Ioup (1988), we gave an explanation of the concept of "sentence intuitions", with examples and practice items, to the participants before the test. These instructions were presented in English for the native English speakers and in Korean for the Korean learners of English.

Subjects

For all three experiments, we had the help of the same participants: a group of 39 native speakers of English and a group of 104 Korean learners of English.

The Korean participants were living in Korea. Seventy of them were students enrolled in the advanced English conversation classes at the language institute of a university in Korea. The remaining 34 were volunteers who met the qualifications of an advertisement posted on the campus: TOEFL 550 or higher, or TOEIC 750 or higher, or 600 or higher on the TEPS. Such non-

native speakers are commonly called "advanced" in second language acquisition studies. None of the subjects had a background in linguistics. (Such background might have exposed them to a discussion of *wanna-*contraction.) The average age of these participants was twenty-six years.

The native-speaker group consisted of people from various places on the U.S. mainland. Nineteen of them were from the West (California or Oregon), ten from the East, six from the Midwest, and four from the South. The average age was twenty-five years. None had training in linguistics.

Order of the experiments

One of us (Kweon) met the participants individually for one-hour appointments. Administration of the three test instruments took about fifty minutes. Each participant was tested on elicited production first, then on the oral repair test, and last on the grammaticality judgement test. Since we wanted to minimize conscious reflection during production and oral repair, we put the grammaticality judgement test last. Of the three, the grammaticality judgement test appears to involve the highest level of metalinguistic awareness. Also, this test included sentences with *wanna* in written form, a fact which might encourage participants to reflect consciously on their use of this form. This ordering of tasks from those involving less reflection to those involving more metalinguistic reflection has become standard practice in second language acquisition research.

Korean-speaking participants were given Korean instructions, followed by English instructions (using the same English instructions given to the native English speakers). Participants were paid a modest fee when they had completed the three experiments.

Common analytical design and rationale

The central goal of our research is to discover whether adult speakers—native speakers and L2 learners—will distinguish between the two types of *wh*-questions, SEQ and OEQ, as exemplified in (1) and (2). Logically, there are four "types" of speakers with respect to the constraints. See Figure 1.

First, it is possible that some will not contract at all; they will contract in neither the theoretically contractible nor uncontractible examples. The

performance of these non-contracters, Group I of Figure 1, does not directly address availability of the constraints, although the possible existence and size of this group is of some interest, given the ready availability of evidence for the contraction process itself.[4]

Figure 1.
Contraction Patterns in OEQ and SEQ in Terms of Obeying the Constraint

	I. conservative	II. correctly differential	III. backward	IV. overgeneral
OEQ	want to	*wanna*	want to	*wanna*
SEQ	want to	want to	**wanna*	**wanna*

   *ungrammatical

The second group, Group II, which we shall call "correctly differential", will contract where it is (theoretically) possible but not when it is impossible. A substantial number of native speakers ought to fall in this group, assuming that the theory is correct and that the experimental instruments are sensitive to the right factors.

The third logically possible group, Group III, which we shall call "backward", would contract where it is impossible but not contract where it is possible. Assuming knowledge of the constraints, no native speaker should be in this category (apart perhaps for some "experimental noise"). Assuming that second language learning is not perverse with respect to input, it is likewise hard to see how learners could fall in this category, except by chance.

Group IV, "overgeneral", would contract in both the possible and the impossible cases.

Expected patterns over the four categories

In the case of native speakers, the "ideal" distribution over these four categories would have everybody in groups I and II: the more in group II the happier for the theory. The grammatical theory would not put native speakers in III or IV. However, the possibility of type-IV "liberal" dialects of native speakers has been raised in the linguistics literature; see especially the discussion in Pullum (1997). We will return to this matter below.

For learners, the interpretation of possible patterns is somewhat less

straightforward. Certainly, an observed distribution that is not very different from the native speaker ideal will tend to support the proposition that the system of abstract empty categories and universal constraints is active in second language acquisition. And, a very large number of speakers in group II—"correctly differential"—would clearly be strong confirmation. In our analysis of results, a key indicator of at least some minimal distinction between the two structural environments is whether there is a statistically significant difference between the proportion of participants falling into category II (correctly differential) and those falling into category III (backward).

At the other extreme, if learners really do not have any idea of the constraint on contraction, so that there is simply a generalized graded tendency to contract (or not), irrespective of structural environment, and if this tendency is spread out at various strengths throughout the population of learners and/or variable within each learner, then learners will fall into all four categories more-or-less at random, depending on how this contraction tendency is distributed. Of course, many complicated intermediate result patterns are possible; these will be discussed below, as evidence for them develops (as it will).

Expected patterns over methods

All three experimental techniques are "performance" measures. That is, they all involve, in ways specific to each test, an interaction between linguistic knowledge and the exigencies of the tasks. Indeed, it is for this reason that multiple methods are advisable, and it is also why the results of each task must be interpreted in light of the particular features (the "distortions") of that task.

In the case of elicited production, a primary concern is the fact that various phonetic processes can influence the pronunciation of consonants: *wanna*-contraction is not the only thing that can cause a stop to be elided or pronounced indistinctly (consider the commonly elided *t* in *Santa Claus*). This could, in principle, result in the "shifting" participants toward category IV.

In the oral repair task—a type of elicited imitation—it is known that speakers, when instructed to repeat, will sometimes repeat even things that their own grammars would not license. (See Bley-Vroman & Chaudron (1994) for an extensive discussion of the strengths and limitations of elicited

imitation.) Since all the stimuli have contraction, a tendency to "parrot" may also shift speakers toward apparent overgeneralization: category IV.

Grammaticality judgements are not so obviously subject to either of these effects, and an extensive comparative study of methods (Masterson, 1993) has shown that judgement tasks can often produce the cleanest results. And, indeed, we expected judgement task to provide the clearest patterns in our study. This is especially so for the native speakers, assuming that the theoretical account is more-or-less correct. Still, there are well-known problems with judgement tasks. They are not "natural", in that making metalinguistic judgements of grammaticality is not something that people do in the ordinary use of language in thought and communication. Subjects may not understand the task correctly. Especially in the case of non-native speakers, there is always the possibility that subjects may be relying on consciously learned rules. In the case at hand, however, there is no evidence that learners would have been presented with any rules relevant to the distinction under investigation.

Categorization criteria

Because the experiments themselves are so different, it is impossible to develop a cross-task scoring system that will allow direct quantitative comparison of the results across the experiments. Nevertheless, we developed a general philosophy of categorization which could provide a certain underlying consistency of rationale in approaching the analysis of all three experiments.

In each of our three experiments, participants are given five or six opportunities to show evidence of contraction, both in SEQs and in OEQs. In principle, if a person showed even one instance of contraction in a given environment, that person could be considered to contract in that environment. However, a preliminary analysis of the data suggested that if a single instance out of five or six were to count as contraction, then almost four fifths of the participants could be said to contract everywhere, even though for many participants (especially in the "impossible" environments), contraction may be quite rare. Thus, a once-is-enough criterion can obscure real patterns. On the other hand, requiring that someone always contract ignores the fact that contraction is, after all, an optional process.

For the elicited production and oral repair tests, we set a cutoff point of more-

than-one, which seems plausible. It ignores the sporadic cases, and when applied consistently seems to give interpretable results, as will be apparent.

For the grammaticality judgement test, we followed the same general philosophy: sporadic cases of acceptance (especially when these are only "probably" acceptable") in an environment will not categorize a participant as a contracter in that environment if there is otherwise clear rejection. At the other extreme, requiring that a participant find all examples "absolutely possible" seems too strict. Because we were able to take advantage of a scale in the judgement task, and because we are dealing here only with the "possibility" of use rather than actual use of contraction, the overall philosophy was implemented somewhat differently in the grammaticality judgement experiment than in the other two experiments. We assigned 0 points to a judgement of "absolutely impossible"; 1 to "probably impossible"; 2 to "probably possible"; and 3 to "absolutely possible". We reasoned that an average score of "probably possible" ought to be sufficient for a participant to count as permitting contraction in a given environment. This means that sporadic cases of "impossible" will not necessarily mean that a participant is categorized as not permitting contraction in an environment. Conversely, an occasional case of "possible" where contraction is otherwise judged "impossible" will not cause a participant to count as a contracter in that environment.

We believe that this approach is sensible on its face. Of course, other analyses are possible, see Kweon (2000) for a discussion of alternative criteria and analytical approaches.

Results

Preliminaries

In this section, all results will be presented in tables of the form given in Figure 2. In this grid, contraction in OEQ goes across, and contraction in SEQ goes down. The categories correspond to those discussed above.

Figure 2: Format for the presentation of results

		OEQ	
		non-contracter	contracter
SEQ	non-contracter	category I conservative	category II correctly differential
	contracter	category III backward	category IV overgeneral

For purposes of exposition, we present the grammaticality judgement results first, followed by the oral repair and elicited production results. It should be kept in mind that the order in which the experiments were actually performed was the reverse of this.

In each of the three experiments, there was a very small number of participants for whom we did not have complete data: typically, items were not finished or responses could not be clearly interpreted. In the analysis of a particular experiment, we excluded any participant with missing or uninterpretable data for that experiment. The total number of participants presented in the results tables for each test will thus not always be the full complement of 39 native speakers and 104 learners.

Grammaticality judgement experiment

Native speakers.

The results for the grammaticality judgement task for native speakers is presented in Table 1. Nearly 100% fell into category II. The striking "cleanness" of this result is confirmation of the general correctness of the theory; it also gives us a degree of confidence in both the experimental technique itself and in the analytic approach we have taken. Thus, it provides a good background against which to interpret other results.

Table 1. Native speaker distribution in the GJT (N=37)

category I (conservative)	category II (correctly differential)
0 (0%)	36 (97.30%)
category III (backward)	category IV (overgeneral)
0 (0%)	1 (2.70%)

Results for non-native speakers are presented in Table 2. In contrast to the native speakers, the greatest number of participants fell in category IV: they judge contraction to be possible in both SEQ and OEQ. However, a fairly large number fell into "correctly differential" category II. As discussed above, a key statistical comparison is between categories II and III. Together, these categories comprise those participants who distinguish between the two structures in the contractibility of *wanna*. Of these, more get it "right" (the 34 in category II) than get it "backward" (the 14 in category III). This difference is statistically significant (z = 2.89, p<.05, one-tailed).[5]

Table 2. Non-native speaker distribution in the GJT (N=98)

category I (conservative)	category II (correctly differential)
8 (8.16%)	34 (34.69%)
category III (backward)	category IV (overgeneral)
14 (14.29%)	42 (42.86%

Oral repair experiment

The native-speaker results for oral repair are shown in Table 3. In contrast to the native-speaker judgement results of Table 1, there are quite a few native speakers "overgeneral" category (IV), approximately the same number as in the "correctly differential" category (III). As in the judgement test, not a single native speaker is "backward".

Table 3. Native speaker distribution in the oral repair test (N=39)

category I (conservative) 5 (12.82%)	category II (correctly differential) 18 (46.15%)
category III (backward) 0 (0%)	category IV (overgeneral) 16 (41.03%)

The results for learners are given in Table 4.

Table 4. L2 learners distribution in oral repair test (N=102)

category I (conservative) 17 (16.67%)	category II (correctly differential) 11 (10.78%)
category III (backward) 2 (1.96%)	category IV (overgeneral) 72 (70.59 %)

For learners, category IV again comprises by far the largest group, with about 70% of the participants. These are learners who show willingness to repeat *wanna* (contracted) in both SEQ and OEQ. A few (about 17%) resist saying *wanna* in either environment (category I). In the key comparison of categories II and III, here too, more get it right (the 11 participants in category II) than get it backward (the 2 learners in category III). The difference is statistically significant ($z = 2.50$, $p<.05$, one-tailed).

Elicited production experiment

The elicited production experiment gives the experimental participants opportunities actually to produce *wanna* in situations where *wanna* might naturally be used (if possible), and in the absence of any prompt containing *wanna*. It is, in a sense, the most "ecologically valid" of the experiments. The results for native speakers are given in Table 5. Most are in category II.

Table 5. Native speaker distribution in the elicited production test (N=39)

category I (conservative) 8 (20.51%)	category II (correctly differential) 21 (53.85%)
category III (backward) 0 (0%)	category IV (overgeneral) 10 (25.64%)

L2 learner results are presented in Table 6.

Table 6. L2 distribution in the elicited production test (N=104)

category I (conservative) 43 (41.3%)	category II (correctly differential) 16 (15.4%)
category III (backward) 9 (8.7 %)	category IV (overgeneral) 36 (34.6 %)

About three-quarters of the learners did not distinguish SEQ and OEQ. They either produced *wanna* in both structural conditions (36 in category IV) or in neither (43 in category I). While there is no statistically significant difference between the correctly differential category (the 16 participants in category II) and the backward category (the 9 in category III) (z = 1.4, p>.05, one-tailed), the difference is in the right direction.

General discussion

Native speaker knowledge of *wanna* contraction

The native speaker results conform well to theoretical predictions, *modulo* the expected characteristics of the experimental tasks.

We interpret the grammaticality judgement results as strongly confirming the published linguistics literature. Almost 100% are correctly differential. It is consistent with these results to say that the grammars of native speakers license *wanna* contraction in OEQ structures but not in SEQ structures, as is expected under the the UG-based account outlined here.

We interpret the native-speaker oral repair results also as broadly consistent with this grammaticality difference. To be sure, some native-speaker

participants repeat *wanna* as *want to*, even when the contracted form is possible; but contraction is an optional process after all. The complete absence of participants in category III is clearly to be expected, given the proposed grammaticality difference. The existence of a fair number of subjects in the overgeneral category (IV) is somewhat more problematic. But, recall that it is known that in elicited imitation experiments, participants may sometimes imitate ("parrot") models which are not licensed by their own grammars.

In the elicited production experiments, native speakers mostly fell into the correctly differential category, and none were in category III, again consistent with the grammaticality difference that was suggested by the other results. As in the oral repair experiment, the presence of some participants in category I is not a problem here, given the optionality of contraction. However, the relatively large number in category IV cannot be given the same "parroting" account that we proposed for the oral repair experiment. In contrast to the oral repair experiment, there is no *wanna* in the experimental materials of the production experiment for the participants to imitate. Here, all instances of *wanna* are being produced spontaneously by subjects. How can this be?

The possible existence "liberal" *wanna* contraction "dialects" among native speakers was alluded to above. Pullum (1997) argues that there are two processes which can result in the elision of the *t* sounds from *want to*. One is the syntactically constrained rule relating *want to* to *wanna*—what we have been calling *wanna* contraction.[6] The other, secondary source of *wanna* is a low-level phonetic/phonological rule of lenition which applies irrespective of syntactic constraints in some dialects. (Recall our mention of dropping the *t* of *Santa Claus*, above.) In a dialect that had this rule, it would apply to both SEQ and OEQ, as long as the purely phonological conditions are met. Pullum says, "I claim that liberal dialect speakers are those who permit the phonological rule reducing *want to* to *wanna* to apply phrasally. I suspect … that it applies within an intonational phrase in the sense of Selkirk 1984." Pullum notes that there could be, in principle, three dialects: a dialect could have the syntactically constrained rule, or the phonological rule, or both.

Let us re-consider the other native-speaker results in the light of this two-source theory of *wanna*. The existence of native speakers in category IV in the production experiment can be seen as confirmation of Pullum's idea. We continue to claim that these native speakers do have a syntactically

constrained rule of *wanna* contraction. However, at least certain of the native speakers have, in addition, a purely phonological rule reducing *want to* to *wanna*. Since this phonological reduction is insensitive to traces, it readily accounts for the participants in category IV.

This two-source explanation can be extended to the oral repair test, to the cases of category IV which we earlier proposed to account for by the task-artifact of "parroting". There, too, the application of a phrase-level phonological rule, in addition to syntactically constrained *wanna* contraction, would tend to shift subjects into category IV. Assuming that parroting also can occur, the compound effects of parroting and of the phonological reduction would provide a ready explanation for why there are even more subjects in category IV in oral repair than in elicited production (where there is no possibility of a parroting effect because there is nothing to parrot). We do not wish to overemphasize this explanatory advantage, since a comparison of the numerical results of these different experiments is not straightforward.

What about the grammaticality judgement results in light of the possibility of two sources of *wanna*? Recall that in the judgement experiment, there is no evidence of a "liberal dialect" among native speakers. Discounting the single stray individual in category IV, everybody is in the "strict" correctly differential category (II). We propose that the grammaticality judgement experiment primarily taps knowledge of syntactically constrained *wanna* contraction, rather than secondary elision within a phonological phrase. Why should this be the case? The judgement test does not require phonetic production. If phrase-level elision is primarily a phenomenon of lenition in speech production, then a non-production test is less likely to be sensitive to it. In addition, the experimental participants are looking at written forms when judging acceptability. Low-level elisions are not standardly represented in English orthography. (Even if we may drop the *t* when we say *Santa Claus*, we don't write *Sanna Claus*.) A reasonable hypothesis is that the written form *wanna*, which the participants read, involves a level in which the syntactically constrained rule of *wanna* contraction is represented, but not the purely phonological process of within-phrase lenition.

In summary, we interpret the results of the three native-speaker experiments as follows:

Figure 3. Interpretation of native-speaker results

	Difference between OEQ and SEQ	Phrase-level phonological lenition	Parroting artifact
Grammaticality Judgement	YES, via UG	NO, because not oral production	NO, because nothing to parrot
Oral Repair	YES, via UG	YES, oral production	YES, because imitation
Elicited Production	YES, via UG	YES, oral production	NO, because nothing to parrot

Thus, our different experiments tease apart the two sources of *wanna*. The native-speaker results definitely show that speakers do have knowledge of the constraints on *wanna* contraction, a fact consistent with the UG explanation; the results also provide evidence for the correctness of the two-source theory and for the differential sensitivity of the various experiments to sources of contraction.

Learner knowledge of *wanna* contraction

The L2 learner results give a rather different picture. Consider first the grammaticality judgement results. While (virtually) all native speakers accept the grammatical examples and reject the ungrammatical examples, among the learners, about as many accept both as make the right distinction. Recall that the native speaker results seem to show that the grammaticality judgement experiment gives the "purest" picture of knowledge of the syntactic constraints, unsullied by parroting effects or by operation of syntactically unconstrained phrase-level lenition. The pattern of a fairly large number of learners failing to make the distinction is repeated in the two other experiments. While native speakers, too, show a blurrier picture in these two experiments than in grammaticality judgement, in every case the learners show a greater tendency not to distinguish than do the native speakers. Also, in every experiment, there are some learners who fall into the "backward" category III; native speakers are never in this category. In the elicited production experiment, the failure to distinguish is very striking. Over 75% of the learners fall into the non-distinguishing categories I and IV, and among those who do make the distinction, there is no significant difference between those who get it right and those who get it backward.

On the other hand, it would definitely wrong to conclude that learners'

grammars do not distinguish the SEQ and OEQ cases. First, in every experiment, there are always more learners in category II than category III.[7] Second, both native speakers and learners show the greatest shifting toward category IV in the oral repair experiment, suggesting that the compound effect of phonological lenition and parroting may be active for learners as well as for native speakers. Third, both learners and native speakers show the greatest tendency toward conservatism (category I) in the oral production experiment. Taken as a whole, these patterns suggest that the experimental tasks are interacting in similar ways with underlying systems which are similar, though not identical. The learners evidently do have at least some grammatical knowledge that has the effect of making contraction less likely in SEQ than in OEQ. This knowledge, however, is much less "certain" than in native speakers. Let us say, tentatively, that learners are indeed sensitive to the syntactic difference, but their knowledge is "imperfect" relative to native speaker knowledge. The concept of imperfect knowledge is intended to describe the general "better-than-chance" performance of learners. It also explains why learners sometimes get it backward (by perverse luck, given the imperfection of knowledge). Additionally, we propose that learners, like native speakers, also have a phonological lenition rule which applies indifferent to SEQ/OEQ difference. This chiefly affects, as is natural, the oral tests: oral repair and elicited production. And, the learner results are also consistent with a parroting effect which boosts the rate of *wanna* production in oral repair, an imitation task. In summary, our account of the L2 learners is shown in Figure 4.

Figure 4. Partial interpretation of learner results

	Difference between SEQ and OEQ	Phrase-level phonological lenition	Parroting artifact
Grammaticality Judgement	YES, but imperfect Why?	NO, because not oral production	NO, because nothing to parrot
Oral Repair	YES, but imperfect Why?	YES, oral production	YES, because imitation
Elicited Production	YES, but imperfect Why?	YES, oral production	NO, because nothing to parrot

Let us consider in greater detail the better-than-chance performance of the learner group, trying to answer the "Why?" of Figure 4. At first glance, the distinction between SEQ and OEQ might suggest that learners' competence must have access to a UG-like system of abstract entities and innate

constraints since they do show some evidence of the distinction between SEQ and OEQ. Certainly, given such a UG system, it is to be expected that learners will develop knowledge of the constraints on *wanna* contraction. All that is required is that learners know two things: First, that the verb want can take infinitival clauses as complements, both with overt subjects and PRO subjects. This much is clear from the readily available positive input, from examples such as *I want you to wash the dishes* and *I want to visit the Grand Canyon*. Second, they must know how to make interrogatives by fronting *wh*-words—something also readily attested in the input (and always taught). The rest—the existence of various sorts of empty categories in the right places and the possibility that contraction may be blocked across an intervening Case-marked trace—are consequences of Universal Grammar. In fact, it is hard to see how learners could fail to master the constraints fully, given access to UG.

But that's the problem: While the learner results are better than chance, they are only better than chance. Learners are not nearly as successful as native speakers and show patterns that are much messier than one might hope to find, given the inexorable operation of Universal Grammar on readily available input. Many learners seem not to distinguish at all and some seem to get the distinction backward. This kind of "good-but-not-great" result is by now commonplace in studies of UG in second language acquisition, and it has seen again and again since the early studies in this area (e.g. Bley-Vroman et al., 1988 for one early example). One might say that access to UG is "incomplete", or "partial", or "impaired", but these labels are not explanations.

On the development of imperfect knowledge

A central challenge to second language acquisition researchers is the construction of an account of imperfect knowledge in the domain which is controlled by UG constraints in native speakers. In this section we take some speculative steps toward such an account, taking the case of *wanna* contraction as a starting point.

How would it be possible for a learner to acquire a "reasonably good" but still imperfect knowledge of these constraints? UG is "too good". While it readily explains uniform successful acquisition of complex and abstract phenomena on the basis of fragmentary evidence, it cannot readily explain variable and partially successful acquisition.

Suppose, however, that the L2 acquisition system makes central use of a mechanism of noticing the (relatively superficial) configurations of heads and complements observed in the data. That such a mechanism should exist is not really controversial; nor is it incompatible with a larger role for UG. In recent work, Hawkins (2001) argues for a theory of "modulated structure building" in which learners, at least in the early stages of acquisition of a syntactic phenomenon, rely chiefly on the complement-selection properties of heads. For example, he analyzes the early-stage development of negation in L2 French as involving relatively superficial complement-selection statements (*pas* selects VP complements with empty heads, for example). This, rather than verb-raising conditioned by the strength of I, is responsible for the word order of negation with respect to verbs in early L2 French (Hawkins, 2001, p. 111). Since such complement-head statements lack the full generality of a UG-based raising account, Hawkins' modulated structure building can explain how different aspects of the "raising parameter" can show up one-by-one in acquisition, despite the fact that they are linked in linguistic theory. For example, in a complement-selection theory, the position of verbs with respect to negation need not be linked to the position of verbs with respect to adverbials; and, the positioning of negation with respect to finite verbs may be (incorrectly) generalized to non-finite verbs.[8]

Suppose now that (at least some) learners are noting the co-occurrence properties (the complement-selection properties) of the contracted form *wanna* itself, in effect treating *wanna* as a separate lexical item.[9]

A learner who approaches *wanna* in this way observes the following basic patterns:

(3)   a. I *wanna* go the the movies.
        b. I *wanna* see John.
        c. I *want* to go to the movies.
        c. I *want* to see John.
        d. I *want* John to go to the movies.

The learner does not observe:

(4)   a. *I *wanna* to go to the movies.
        b. *I *wanna* to see John.
        c. *I *wanna* John (to) go to the movies.

The learner's grammar can thus develop the following complement-selection rules for *want* and *wanna*:

(5)   a. WANT selects infinitive-clause complements (either with or without an overt subject), such as [*to go to the movies*] or [*John to go to the movies*].
b. WANNA selects bare-infinitive VPs as complements, such as [*see John*] or [*go to the movies*].

In this respect, these learners are treating *wanna* something like a modal verb, though of course they may also notice that it differs from modals in other respects: *wanna* combines with *do* to form questions, for example. Crucially, there is no evidence in the input that *wanna* can select complements of the form [*John go to the movies*].

Two key facts about learning complement-selection are that it is heavily dependent on input and that it is not guaranteed to succeed. In these respects, it differs from the constraints of UG. You cannot know for sure what complements a verb takes without positive evidence. Not all learners will learn all verbal complement patterns. And, a given learner may have imperfectly mastered a particular complement pattern. This is the way it is with lexical learning. Not all learners know that the English word for "tailpipe" is *tailpipe*, and a given learner may be only, say, 75% sure that the word for "tailpipe" is *tailpipe*. Not all L2 learners come to know that *depend* takes *on*-phrases as complements. Not all learn that the location argument of *enter* is expressed as a direct object NP (rather than, say, a prepositional phrase with *into*.) Likewise, not all L2 learners will record the fact that *wanna* only occurs with bare-infinitive VPs, and some learners may be more certain than others about this co-occurrence. Further, some learners may not treat *wanna* as an item with its own complement-selection properties; they may view it simply as a contraction of *want to*.

Now, consider the consequences of this inherently variable knowledge of the complement-selection of *wanna* on the asymmetry between OEQ and SEQ with respect to contraction. The learner knows that *wh*-questions are formed by extracting a *wh*-word and moving it to the front of the sentence. Extraction is possible from within an embedded VP complement, from both subject and object position of the complement, as is clear from potential input examples like those of (6).

(6)   a. Who did you tell Mary to visit *t*?
        b. Who did you see *t* run the marathon?
        c. Who do you want *t* to win the prize?

A learner who has noticed that *wanna* takes bare-infinitive complements cannot create subject-extraction questions with *wanna*. For such a learner, (7a), below, is impossible precisely because (7b) is impossible. Example (7b) is impossible because it violates the complement-selection rule (5b) for *wanna*. And, (7c) is possible because (7d) is possible.

(7)   a. *Who do you wanna go to the movies.
        b. *I wanna John go to the movies.
        c. Who do you want to go to the movies.
        d. I want John to go to the movies.

On the other hand, a learner who has not learned the complement-selection rule, or who has learned it imperfectly, and who has instead, or in addition, a low-level rule phonological contracting *want to* to *wanna*, can produce (7a) simply by contraction from (7c).

Figure 5. Interpretation of learner results

	Difference between OEQ and SEQ	Phrase-level phonological lenition	Parroting artifact
Grammaticality Judgement	YES, but imperfect, via complement selection	NO, because not oral production	NO, because nothing to parrot
Oral Repair	YES, but imperfect, via complement selection	YES, oral production	YES, because imitation
Elicited Production	YES, but imperfect, via complement selection	YES, oral production	NO, because nothing to parrot

In summary, imperfect knowledge of the constraints on *wanna* contraction can be given an account within a general theory of adult SLA which has a central role for the learning of complement-selection properties of heads, combined with the possibility that at least some learners will consider *wanna*

a head with its own complement-selection properties. Variability is expected on this account, on the assumption that different learners notice different things, and that a given learner can have degrees of certainty (or "mastery") over a given lexical property.[10] Learners with this sort of imperfect knowledge can, in principle, exhibit behavior which ranges from a level that is barely better than chance to a level which is indistinguishable from the "perfection" of a categorical constraint. Figure 5 answers the "Why?" of Figure 4.

References

Ayoun, D. (to appear, 2002). <u>Parameter-setting theory in first and second language acquisition</u>: Continuum.

Bergen, B. K. (2000). Probability in phonological generalizations: modeling French optional final consonants. <u>BLS</u>.

Bley-Vroman, R. (1997). <u>Features and patterns in second language acquisition.</u> Paper presented at the Second Language Research Forum, Michigan State University, East Lansing, Michigan.

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. Tarone & S. Gass & A. Cohen (Eds.), <u>Research Methodology in Second-Language Acquisition</u> (pp. 245–262). Hillsdale, N.J.: Lawrence Erlbaum.

Bley-Vroman, R., Felix, S. W., & Ioup, G. (1988). The accessibility of Universal Grammar in adult language learning. <u>Second Language Research, 4</u>(1), 1–32.

Bolinger, D. L. (1980). Wanna and the gradience of auxiliaries. In G. Brettschneider & C. Lehmann (Eds.), <u>Wege zur Universalienforschung</u> (pp. 292-299). Tübingen: Günter Narr.

Bolinger, D. L. (1981). Consonance, dissonance, and grammaticality: the Case of wanna. <u>Language and Communication, 1</u>, 189-206.

Chomsky, N. (1976). Conditions on rules of grammar. <u>Linguistic Analysis, 2</u>, 303–351.

Chomsky, N. (1977). On wh-movement. In P. Culicover & T. Wasow & A. Akmajian (Eds.), <u>Formal syntax</u> (pp. 71-132). New York: Academic Press.

Chomsky, N. (1980). <u>Rules and Representations</u>. New York: Columbia University Press.

Chomsky, N., & Lasnik, H. (1977). Filters and control. <u>Linguistic Inquiry, 8</u>, 425-504.

Chomsky, N., & Lasnik, H. (1978). A remark on contraction. <u>Linguistic Inquiry, 9</u>, 268-274.

Crain, S., & Thornton. (1998). <u>Investigating Universal Grammar</u>. Cambridge, Mass.: MIT Press.

Eubank, L. (Ed.). (1991). <u>Point Counterpoint: Universal Grammar in the Second Language</u>. Amsterdam: John Benjamins.

Fiengo, R. (1980). <u>Surface Structure</u>. Cambridge, Mass.: Harvard University Press.

Hanley, J. A., & Lippman-Hand, A. (1983). If nothing goes wrong, is everything all right? Interpreting zero numerators. Journal of the American Medical Association, 249(13), 1743-1745.

Hawkins, R. (2001). Second Language Syntax: A Generative Introduction. Oxford: Blackwell.

Hawkins, R., & Chan, Y.-h. (1997). The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. Second Language Research, 13(3), 187–226.

Herschensohn, J. R. (2000). The second time around : minimalism and L2 acquisition. Amsterdam ; Philadelphia: J. Benjamins Pub.

Jaeggli, O. (1980). Remarks on to contraction. Linguistic Inquiry, 11(239-246).

Karins, A. K., & Nagy, N. (1993). Developing an experimental basis for determining grammaticality. Penn Review of Linguistics, 17, 93-100.

Kellerman, E. (1977). Toward a characterisation of the strategy of transfer. Interlanguage Studies Bulletin, 2, 59–92.

Kweon, S.-O. (2000). The acquisition of English contraction phenomena by advanced Korean learners of English: Experimental studies on wanna contraction and auxiliary contraction. Unpublished PhD dissertation, University of Hawai'i, Honolulu.

Kweon, S.-O. (2001a). UG Availability and Wanna Contraction. Studies in Modern Grammar, 24, 203-223.

Kweon, S.-O. (2001b). The acquisition of contraction constraint by non-native speakers of English. English Teaching, 56(3), 89-109.

Lakoff, G. (1970). Global rules. Language, 46, 627-639.

Lightfoot, D. (1976). Trace theory and twice-moved NPs. Linguistic Inquiry, 7, 559-582.

Lightfoot, D. (1977). On traces and conditions on rules. In P. W. Culicover & T. Wasow & A. Akmajian (Eds.), Formal Syntax (pp. 207-237). New York: Academic Press.

Masterson, D. (1993). A Comparison of Grammaticality Evaluation Measurements: Testing Native Speakers of English and Korean. Unpublished PhD, University of Hawai'i.

Munnich, E., Flynn, S., & Martohardjono, G. (1994). Elicited imitation and grammaticality judgement task: What they measure and how they relate to each other. In E. Tarone & S. Gass & A. Cohen (Eds.), Research Methodology in Second-Language Acquisition (pp. 227-243). Hillsdale, N.J.: Lawrence Erlbaum.

Pullum, G. K. (1997). The morphological nature of English to-contraction. Language, 73, 79-102.

Rotenberg, J. (1978). <u>The Syntax of Phonology</u>. Unpublished Doctoral dissertation, MIT, Cambridge, Mass.

Selkirk, E. O. (1984). <u>Phonology and Syntax</u>. Cambridge, Mass.: MIT Press. (Cited in Pullum (1997).

Notes

[1] In Karin and Nagy's study, the participants were given a scenario in which an old woman and a man are struggling over the woman's purse. The woman cries for help. The prompt concludes as follows: "A policewoman and a big strong guy are standing by. Which one would you want to [wanna] help?" Given that the immediate prior context has set up the alternatives of a "policewoman" and a "big guy", it is likely that "which one" will be interpreted as referring to one of these--a subject interpretation. As Karin and Nagy put it, the context is "pragmatically skewed." One standard way of understanding the results is that grammatical constraints can interact with pragmatic/discourse effects in performance. Given the right context, pragmatics can override grammar. Frequently, we are able to understand what people must have meant even though it's not precisely what they have said. This is hardly controversial. Still, the mechanisms of the interaction remain very poorly understood, and Karin and Nagy do not to favor such an approach. They point out that explanations which involve the interaction of separate systems of syntax and pragmatics are difficult to investigate empirically, and they suggest an alternative in which one simply says, directly, that *wanna*-contraction is a "variable rule."

[2] The present paper is based on the results of Kweon (2000); some of these results are also reported in Kweon (2001a, 2001b). Kweon (2000) reports the details of the experiments and several alternative statistical analyses. The interpretations of the present work are substantially different from and go beyond those of that earlier work. We wish to thank the East-West Center in Honolulu for the field research grant which made possible our data collection.

[3] Also included in the grammaticality judgement test, but not analyzed in the results here, were five each SEQ and OEQ occurring as embedded *wh*-interrogatives in declarative sentences as complements of verbs taking indirect questions, e.g. *I know what they wanna read in the newspaper* or *I know who the boys wanna make pizza*. These types were originally included to explore a separate question: the interaction of the choice of main-clause sentence type and complexity (declarative or interrogative, reflecting a difference in the number of nodes crossed by the *wh*-word) with the *wanna* -contraction asymmetry. In fact, we found no significant effects on this "fishing expedition." It doesn't matter whether the extracted *wh*-word is at the front of a main-clause interrogative or embedded in an indirect question. We exclude these embedded examples in order to maintain comparability of sentence type across all experiments and so to facilitate comparison between the experimental results.

[4] There are over 28,000 examples of the sequence *want to* in the approximately 6,000,000 sentences of the British National Corpus: about five examples per one thousand sentences. The examples seem, on a first inspection, to be especially dense in spoken language or in genres where there is substantial amounts of represented conversation, e.g. fiction: "I want to [wanna] see you on your knees, begging." One cannot be sure, of course, without exhaustive study of all 28,000, how many of these examples will have contraction blocked by

an intervening trace. A hand-analysis of 1,000 randomly selected examples found none in which contraction would have been blocked, giving an estimated upper limit of non-contractible examples at about 0.3% of the total (using the Hanley/Lippman-Hand (1983) approximate method for estimating a 95% confidence interval with zero occurrences). In any case, the corpus data seem to show (a) that the opportunities for *wanna* are very numerous and (b) that the cases where contraction would be blocked are very few.

[5] Asking whether more people get it right than get it backwards constitutes a reasonable minimal statistical test: a foundation on which to build a more nuanced qualitative discussion of result patterns. We say this test is "minimal", since, after all, the theoretical claims of Universal Grammar are not merely that something obtains "more often than not." So, we do not immediately say, when we see a "more-do-than-don't" result, that Universal Grammar guides second language acquisition. Of greater importance is whether a coherent account can be offered for the complex pattern of the results. Furthermore, when we say that the claims of UG are "not merely statistical" we do not mean to say that statistics is irrelevant to studies of UG. We do consider tests of statistical significance to be an essential component. They are necessary, but not sufficient to confirm a UG-based theory.

There are, obviously, several other approaches to a minimal test of significance. One could, for example, consider categories I and II to be "constraint-compatible", while III and IV are "constraint-incompatible." Then, one might ask if there were significantly more "constraint compatible" participants than "constraint-incompatible" participants. The native speakers would always "pass" this test, while the learners never would.

Another, more often used approach, is simply to perform an analysis of variance and ask whether there is a significant effect for the syntactic factor in question. If this route is taken, in every experiment, for both native speakers and learners, there is a significant effect on contraction of the difference between OEQ and SEQ. Again, this is only a first step.

A slightly different question is whether the distribution into categories is significantly different from the distribution which would be expected if everyone were guessing (more precisely, if the responses throughout the whole collection of response data for a given experiment were random). For what it's worth, all the results for all experiments, for both the natives and the non-natives, are significantly different from such a random distribution. For a discussion of these different statistical approaches, see Kweon (2001a).

[6] Pullum's account uses syntactic subcategorization rather than blocking by traces to constrain *wanna* contraction. A separate lexical item *wanna* is related to *want* by a morpholexical rule. There are, of course, important differences between a morpholexical account and a trace-theoretic account. Pullum's essential insight, for our purposes, is his distinction between syntactically constrained contraction and purely phonological contraction. At this level, it does not matter whether the syntactic constraints are handled by subcategorization or by trace theory.

[7] Even in oral production, the difference is close to being significant. Just a few more

participants in the learner pool would make it so (assuming they showed the same pattern). The fact that the other experiments do show a significant effect also suggests that it might be an error to accept the null hypothesis in the case of elicited production. Alternative analytical approaches can show a significant difference. For example, if one ignores the categorization altogether, an analysis of variance shows a significant effect for syntactic structure on learner production of *wanna* even in the production experiment.

[8] The failure of phenomena united in a single UG parameter to be acquired in a unitary way is well-established in L2 acquisition studies. For excellent overviews of especially the Romance data, see Herschensohn (2000) and Ayoun (to appear). Both Ayoun and Herschensohn discuss the difficulties of "piecemeal" acquisition for UG-based SLA theory, and they make some suggestions of how these difficulties are to be overcome. The concept of "construction-by-construction" acquisition of Bley-Vroman (1997) seems to be in the same vein as Hawkins' proposals, though couched in a rather different theoretical framework. Bley-Vroman, in contrast to others, does not propose that learners eventually make the transition from construction-based statements to UG-like parameters. The evidence of our present experiment does not address this issue.

[9] That *wanna* can be treated as a word in its own right has also proposed by other researchers in English grammar; we do not claim originality for this basic idea. Dwight Bolinger, notably, suggested that *wanna* had many, though not all, of the properties of an auxiliary verb (Bolinger, 1980, 1981). Pullum, also, whose two-source account of *wanna* was invoked in the explanation of our results above, has argued that *wanna* could be treated as a separate lexical item.

[10] Building degree of certainty into a model of L2 knowledge might go a ways toward accounting for variable learner performance. One approach to modeling might be Bayesian statistics, Bayesian belief networks, in particular. In a learner's grammar, principles might have probabilities associated with them—these probabilities interpreted as certainty levels and derived in part from input frequency and in part from expectations (interpreted as Bayesian prior probabilities) based on native language characteristics, and possibly other factors such as "transferability" (Kellerman, 1977). In pioneering work, Bergen (2000) employs Bayesian belief networks in an account of variable performance of native speakers of French on consonant liaison. Bergen does not interpret his probabilities as certainty, but rather as the speaker's knowledge of actual probabilities. The mathematics, however, are the same.