

Want-to* Contraction in Second Language Acquisition: An Emergentist Approach

William O’Grady, Michiko Nakamura & Yasuko Ito

1. Background

The standard view of the computational system for language is that it incorporates a grammar—a set of ‘rules’ (constraints, principles) that specify the properties that a well formed sentence must have. This is the view adopted in essentially all formal work on language, including of course theories that take Universal Grammar to be at the heart of the human language faculty.

Standing in contrast to this view is the emergentist thesis, which holds that the properties of language are shaped by more basic, nonlinguistic forces (e.g., Elman 1999, MacWhinney 1999, Menn 2000, O’Grady this issue). A particular version of this thesis is developed in some detail by O’Grady (2005), who proposes that the core properties of natural language syntax follow from the operation of an efficiency-driven computational system that is indistinguishable from a processor. (See O’Grady 2001 for a summary.)

The computational system outlined by O’Grady does exactly what any processor does: it operates in a linear manner, it combines elements, and it checks to make sure that lexical requirements (‘dependencies’) are being satisfied. But unlike conventional processors, it is not constrained by grammatical principles. Rather it simply seeks to reduce the burden on working memory by carrying out its operations at the first opportunity (the ‘Efficiency Requirement’).

As a simple example, consider the formation of the sentence *Jerry visited Boston*, which is built around a verb that requires two nominal arguments.

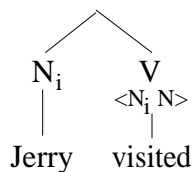
(1) *visit*: V, <N N>

The computational system first combines *Jerry* and *visited*, creating the structure depicted below and resolving the verb’s first argument dependency. (We indicate that a dependency has been resolved by copying the index of the nominal into the appropriate position in the verb’s grid. The order in which dependencies are

* We are very grateful to Kevin Gregg, Soo-Ok Kweon, Dean Mellow, Yuko Otsuka, Amy Schafer, Bonnie Schwartz, and two anonymous referees for their helpful comments and suggestions.

resolved is determined by their position in the verb's argument grid, starting with the leftmost dependency.)

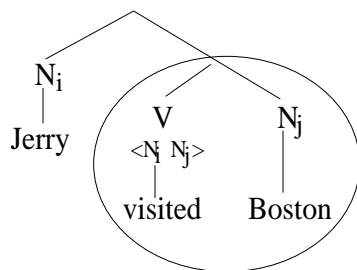
- (2) Combination of *Jerry* and *visited*; resolution of the verb's first argument dependency



in abbreviated form: [Jerry_i visited]
<N_i N>

The computational system next combines *visit* with the nominal *Boston*, creating an opportunity to resolve the verb's second argument dependency.

- (3) Combination of *visited* and *Boston*; resolution of the verb's second argument dependency



in abbreviated form: [Jerry_i [visited Boston_j]]
<N_i N_j>

As illustrated here, the computational system consistently does what is required of it without delay. Thus, if there is an opportunity to combine two elements, as happens upon encountering the words *Jerry* and *visited* at the beginning of the sentence, that opportunity is immediately exploited. And if this then creates an opportunity to resolve an argument dependency, as it does here, this too is done immediately. And so on, until the sentence is fully formed.

2. The acquisition problem

A major advantage of UG-based work on language is its commitment to the question of how children can acquire the intricacies of human language in response to the particular sorts of experience to which they have access.

The intellectual interest of this problem stems largely from the fact that a number of ‘core’ linguistic phenomena exhibit properties whose abstractness and complexity appear to rule out induction from experience, especially the type of experience to which children have access during the crucial early years of life. One such phenomenon involves the constraint on contraction whose effects are observed in the following familiar contrast.

- (4) a. (Guess) who they want to/wanna see.
 (cf. They want to see who.)
- b. (Guess) who they want to/*wanna stay.
 (cf. They want who to stay.)

The commonly held view in the Principles-and-Parameters tradition is that contraction in the second sentence is blocked by the trace of *wh* movement in the embedded subject position.

- blocks contraction
 ↓
- (5) (Guess) [CP who_i [IP they want [CP t_i [IP t_i to stay]]]]. (= (4b))
 [+Case]

As the acceptability of (4a) demonstrates, however, contraction is not blocked by the intermediate trace in the specifier of CP or by PRO.

- do not block contraction
 ↓ ↓
- (6) (Guess) [CP who_i [IP they want [CP t_i [IP PRO to see t_i]]]]. (= (4a))

Of the many proposals that have been put forward over the years to account for this contrast, perhaps the best known is Jaeggli’s (1980) idea that only Case-marked traces block contraction. Thus, contraction is prohibited in (5), where the trace to the immediate left of *to* carries objective Case, but not in (6) where the two elements intervening between *want* and *to* lack Case.

The consequences of Jaeggli’s proposal and of its descendents (e.g., Barss 1995, Boeckx 2000) extend well beyond the particular facts of *want-to* contraction. If such analyses are on the right track at all, the contrast for which they account can hardly be learned from experience. Generalizations involving invisible traces—let alone Case-marked invisible traces—are not particularly good candidates for inductive learning. As White (1989:7) puts it:

None of [the necessary] information is obviously present in the input, since traces are an abstraction. The fact that *wh*-movement leaves a trace and that this trace [when Case-marked] blocks the operation of certain rules is knowledge derived from Universal Grammar, and not from the input alone, or from any general non-linguistic cognitive principles.

See Crain & Thornton (1998:23) for additional remarks along these lines.

We agree that constraints on contraction and other abstract syntactic phenomena (coreference, control, agreement, extraction, and so forth) are unlikely to be induced from experience. The complexity of the facts, the limited availability of relevant experience, the speed of acquisition, and the general absence of mistakes all belie the sort of trial-and-error learning associated with induction.

On the other hand, we reject the idea that the properties of syntactic phenomena are given by an innate Universal Grammar. On the view we adopt, Universal Grammar is not needed to account for how language is acquired, any more than it is needed to account for why language has the particular properties that it does. In each case, the burden of explanation falls on the processor—the efficiency-driven linear computational system that we take to lie at the heart of the human language faculty.

The contraction facts offer an opportunity to illustrate this point.

3. An emergentist account of contraction

The particular intuition that we wish to develop is simply that contraction is most natural where the computational system is able to combine the involved elements *immediately*.¹ More precisely:

(7) *Natural Contraction:*

Contraction of the string XY is most natural when X combines with Y without delay.

¹ Although we do not rule out the possibility that *wanna* exists independently as a lexical item (e.g., Pullum 1997), we assume that it can also be formed from *want to* via contraction in the course of actual speech.

This makes good phonetic sense, of course, since the articulatory processes involved in contraction (assimilation, deletion, vowel reduction, and so forth) are most common and most natural in rapid connected speech.

But what precisely counts as a ‘delay’ in a left-to-right computational system? Following O’Grady (2005:139ff), our proposal is as follows:

- (8) In the pattern ABC, A and B combine without delay if they combine immediately. B and C combine without delay if they combine right after the computational system has finished with A and B.²

A simple example of the first case involves *She will go*, where *she* is A, *will* is B, and *go* is C. Proceeding in the usual left-to-right manner, the processor combines *she* and *will* as soon as they are encountered, thereby paving the way for contraction to *she’ll*.

- (9) Immediate combination of the subject with the auxiliary verb, creating an opportunity for contraction:

contraction gives *she’ll*
↓
[she will]

A simple example of the second case involves *They could have (left)*, where *they* is A, *could* is B, and *have* is C.

- (10) They could have left.

Step 1: Combination of *they* and *could*.

[they could]

Step 2: Combination of *could* with *have*:

contraction gives *coulda*
↓
[they [could have]]

The key step here is the second one. Immediately after the computational system has finished with *they* and *could*, it combines *could* and *have*—creating the conditions that invite contraction.

² That is, as soon as the computational system has combined A with B, and, where appropriate, used one to resolve an argument requirement of the other.

3.2 Two uses of *want*

A second prerequisite for understanding contraction involves the fact that *want* occurs in two quite different ‘frames’ in non-question structures.

- (13)a. They want to see Rick.
 b. They want Mary to stay.

We assume that *want* takes two arguments in the first pattern—a nominal to the left and the infinitival marker *to* to its right. (The infinitival marker in turn takes a verbal argument.)

- (14) *want*₁: <N TO>

The status of *want* in (13b) is somewhat more controversial, but we side with Gazdar, Klein, Pullum, & Sag (1985:145) in assuming that it takes two nominal arguments in addition to an infinitival argument.

- (15) *want*₂: <N N TO>

This contradicts the view, common in Government and Binding theory, that the second nominal in (13b) functions as subject of the embedded verb, as depicted below.

- (16) They want [Mary to stay].

Independent evidence for the view we adopt comes from Postal’s (1974:186) observation that a sentential adverb with scope over the matrix clause can appear to the right of the postverbal NP.

- (17) They wanted Mary, *unwisely*, to serve as their president.
 (= ‘They unwisely wanted Mary to serve as their president.’)

If *Mary* is in the embedded clause, the adverb must be as well, contrary to its interpretation as modifier of *want*. On the other hand, if *Mary* is part of the matrix clause as we propose, *unwisely* can be too, in accord with its semantic function.

Further evidence for the same conclusion comes from the acceptability of the negative polarity item *any* in sentences such as the following.

- (18) I want *none of the money* to go to politicians, [*any* more than you do].
 (cf. *I want some of the money to go to politicians, any more than you do.)

Given that negative polarity items must be licensed by a structurally higher negative in English and that the *any* phrase in (18) clearly belongs in the matrix clause (*any more* modifies *want*), it follows that *none of the money* must be in the top clause as well—just as we’d expect if it serves as an argument of *want*.

With these background assumptions in place, let us turn to the question of why the <N TO> *want* pattern but not its <N N TO> counterpart, permits fully natural contraction in *wh* questions.

- (19)a. The <N TO> pattern
 (Guess) who they want to/wanna see.
 (cf. They want to see Rick)
- b. The <N N TO>
 (Guess) who they want to/*wanna stay.
 (cf. They want Mary to stay.)

Our goal will be to show that the contrast follows from the fact that contraction is most natural when the elements involved combine without delay, as defined in (8) above, repeated here as (20).

- (20) In the pattern ABC, A and B combine without delay if they combine immediately. B and C combine without delay if they combine right after the computational system has finished with A and B.

3.3 The permissible pattern of *want-to* contraction

The formation of a *wh* question such as (19a), which involves the <N TO> use of *want*, is summarized below.³ (We postpone until section 6 the crucial question

³ In the theory outlined by O’Grady (2005), the computational system has a mandate to structure the input as quickly as possible and therefore temporarily combines the *wh* word and the subject, even though there is no functor-argument relation between them. Phonological effects, such as the flapping of /t/ in ... *wha[D] it did*, provide one indication that the two words do in fact merge in the course of production. For further discussion, see O’Grady (2005:21ff & 114); Steedman (2000) outlines a different sort of computational system that also allows immediate combination.

depicted in step 3, it takes advantage of the opportunity to resolve the *wh* dependency associated with *want*'s second argument—the *wh* word *who*.

From the point of view of both comprehension and production, contraction in (23) therefore creates a contradiction. Whereas *wanna* signals that *want* has combined with *to* without delay, this is not what happens. Rather, right after the computational system finishes with *they* and *want*, an opportunity arises to resolve the *wh* dependency by linking it to the second argument slot in the verb's grid. Given the Efficiency Requirement, this opportunity must be exploited, thereby delaying combination of *want* with *to* (its third argument).

This contrasts with the situation found in patterns such as *Who do you wanna see?*, where contraction is permitted. There, as we saw in (21), the computational system combines *want* with *to* immediately after dealing with *they* and *want*.

The difference between the two *want to* patterns thus comes down to the timing of the contraction operation. Where *want* and *to* combine immediately, contraction is permitted; where there is a delay, contraction is less than fully natural.

There appear to be prosodic correlates of this contrast. Warren, Speer, & Schafer (2003) report that *want* has a consistently longer duration in patterns where contraction is blocked. In addition, patterns of this type are more likely to manifest an intermediate prosodic break after *want* than are patterns in which contraction is permissible. It is not implausible to think that both the duration of *want* and the appearance of a prosodic break right after it are reflexes of the time required to resolve the *wh* dependency—time that ultimately compromises the naturalness of contraction.

4. The generality of the *want to-wanna* contrast

Although there is, as we have suggested, a principled basis in processing for the *want to/wanna* contrast, it is apparently not exploited by all native speakers of English. On the one hand, there are varieties of English that simply do not permit *want to* contraction at all—e.g., British English, according to Bailes (2000). On the other hand, there have long been anecdotal accounts of speakers who accept and produce *want to* contraction even in sentences such as (24) (e.g., Postal & Pullum 1982, Pullum 1997).

(24) (*)Who do you wanna stay?

‘Liberal’ speakers have also been identified in at least one experimental study—Ito (2005) reports that 5 of her 41 native English subjects accepted patterns such as (24).

It is relatively easy to accommodate the existence of such speakers in the sort of approach we adopt. That is because our constraint, repeated below, neither requires nor categorically prohibits contraction.

(25) *Natural Contraction:*

Contraction of the string XY is most natural when X combines with Y without delay.

Rather, it simply predicts that contraction will be less natural (i.e., more marked) in patterns such as (24). Thus most speakers avoid *wanna* in such cases and find it unnatural in the speech of others. But nothing in principle rules out the existence of varieties of speech that permit the marked case.

The study of relative clauses provides an instructive precedent. From a processing perspective, subject relative clauses appear to be the least demanding of all relative clause types to produce and comprehend (see O’Grady 1997:178ff and Hawkins 1999, 2004:169ff for extensive discussion). And, indeed, there are some languages that permit only subject relative clauses (e.g., Keenan & Comrie 1977). However, there are of course also languages that permit both subject relatives and direct object relatives (as well as languages that have an even wider range of relative clause types).

Facts such as these do not in any way undermine the processing explanation for markedness asymmetries. Rather, as Hawkins (2004:256ff) notes, the prediction is simply that if the harder pattern is possible, then so must be the easier pattern. Hence if a language permits direct object relatives, it must also allow subject relatives—which does in fact seem to be the case.

A parallel prediction can be made for the *want to/wanna* contrast. In particular, we predict that any variety of English speech that permits non-immediate contraction in patterns such as (24) must also permit immediate contraction in patterns such as (26).

(26) Who do you wanna see?

What we predict to be impossible is a variety of speech that permits contraction in (24), but not in (26). As far as we know, no native speaker of English exhibits this pattern of judgments.

5. *Want-to* contraction in second language acquisition

If the analysis just outlined is correct, then the facts of *want-to* contraction follow straightforwardly from the interaction of an efficiency-driven processor with the human articulatory apparatus—for the reasons noted earlier, the phonetic operations responsible for contraction are most natural when combination is immediate. There is thus no need either for an innate grammatical constraint on contraction or for an induction-based account of how the contrast is learned.

This proposal fits well with what is known about first language acquisition, where the contrast appears to emerge spontaneously and with few or no errors. Indeed, experimental work by Crain & Thornton (1998:184) suggests that children as young as two distinguish between the contractible and noncontractible patterns with a high degree of accuracy. Curiously, however, matters are quite different in the case of second language acquisition.

Pioneering work by Kweon (2000) suggests that second language learners encounter significant difficulty with the phenomenon of *want-to* contraction. Kweon gave 98 adult Korean native speakers (all advanced EFL learners) a battery of tasks related to contraction, including a grammaticality judgment test. When asked to judge the status of sentences such as *I know who they wanna clean the table*, which were simultaneously presented in spoken and written form, 79 of Kweon's subjects (80%) consistently accepted them.

Roughly comparable results are reported by Ito (2005), who elicited grammaticality judgments from 93 Japanese-speaking learners of English as a foreign language (30 at the beginning level, 30 at the intermediate level, and 33 at the advanced level.) Ito's test sentences were presented in contexts such as the following.

There will be a job fair next week, and Audrey wants someone to interview Susan at the fair. David is curious to know who it will be, and asks Audrey:

Who do you wanna interview Susan at the job fair next week?

The subjects were asked to rate the sentences as perfect, okay, strange, or awful (or to indicate if they had no opinion.) In addition, they were asked to rewrite any sentence that they judged to be strange or awful, or about which they had no opinion.

Ito converted subjects' judgments to numerical scores (4 for 'perfect', 3 for 'okay', and so on), choosing a mean score of 3.2 out of 4 (80%) as the cut-off for

distinguishing between acceptance and rejection of the two contraction patterns. Table 1 summarizes the performance of her subjects in this regard. Conservative subjects rejected both patterns, and overgeneralizers accepted both.

Table 1 Performance of subjects in Ito's experiment

Level	Conservative	Overgeneralizers	Native-like	Other
Beginning	8	14	6	2
Intermediate	9	12	8	1
Advanced	9	9	13	2

As can be seen here, almost half of Ito's beginning subjects (14 out of 30) accepted the forbidden contraction pattern. The proportion of intermediate and advanced students who overgeneralized is somewhat smaller, but is still substantial.

Evidently then, even some advanced learners of English have trouble with the phenomenon of *want-to* contraction. Why should this be?

6. Explaining the L2 facts

The particular problem with which we are confronted is basically this: a large number of second language learners behave as if they were 'liberal' speakers of English, permitting contraction in patterns such as (27), where more conservative native speakers forbid it.

(27) *Who do you wanna stay?

This in turn leads to the following research question, which we take to be the central issue in the study of *want to* contraction in second language acquisition.

(28) Why do NNSs exhibit a greater propensity than NSs to be liberal in their acceptance of *want to* contraction?

As we have already noted, the constraint that we have proposed does not categorically rule out contraction in these cases. It simply presents a processing-based dividing line (immediacy) such that one type of contraction is more 'natural' (i.e., less marked) than the other. But what factors might lead second language learners to ignore this dividing line?

In the absence of Universal Grammar, there are only two places to look for an explanation—in the functioning of the processor and in the properties of the lexical items to which it applies. The former possibility has in fact been considered for a variety of filler-gap phenomena, leading to the suggestion that L2 learners do not retrieve or reactivate fronted *wh* words at structurally defined ‘gap’ sites. Instead, it is suggested, they rely on lexico-semantic information to conduct a ‘shallow parse’ that ignores many syntactic details (e.g., Marinis et al. 2005, Clahsen & Felser 2006a,b, Felser & Roberts 2007).

If the *wh* word is not reactivated at the ‘gap site’ in patterns such as *Who do you want _ to stay?*, resolution of the *wh* dependency in these sentences might not interfere with immediate combination of *want* and *to*. This in turn could remove the impediment to contraction and nullify the *want to/wanna* contrast. At the same time, the contrast might collapse in a different direction if the L2 processor were simply slow, as is frequently acknowledged (e.g., Williams, Mobius & Kim 2001, and Clahsen & Felser 2006b and the references cited there). Slowness might lead to underexploitation of the opportunity to contract, even in acceptable patterns such as *Who do you want to see?* Interestingly, Kweon (2000:170) reports that 60 of her 104 adult Korean-speaking (mostly advanced-level) subjects failed to contract *want* and *to* even where it was permissible to do so in the elicited production task that she conducted.

While we do not rule out these sorts of explanations, we wish to consider another (perhaps complementary) possibility involving the lexical properties of *want* itself.

6.1 Lexical properties

In considering the role of lexical properties in the operation of the processor, it is important to distinguish between the propensity to resolve dependencies at the first opportunity and the identification of those opportunities. Whereas the former is an internal feature of the computational system (a manifestation of its drive to reduce the burden on working memory), the latter is potentially sensitive to external factors such as relative frequency of occurrence, pragmatics, and the like.

It has long been recognized that holding a *wh* dependency in working memory creates a burden that is best discharged as quickly as possible. According to the Active Filler Hypothesis (e.g., Clifton & Frazier 1989), for instance, the processor actively seeks out the first potential ‘gap’ to which it can link the *wh* dependency (see also Crain & Fodor 1985, and De Vincenzi 1991).

(29) The Active Filler Hypothesis

Assign a filler to a gap (= ‘resolve a *wh* dependency’) as quickly as possible.

At the same time, though, it is well known that ‘active filler’ effects can be mitigated by lexical considerations. We are therefore less troubled by the late association of the *wh* filler with a gap in (30a), where the verb is usually intransitive, than in (30b), where it is usually transitive (Fodor 1978).

(30)a. Who did you walk () to the store with _?

b. Which book did the teacher read () to the students from _?

We propose that external factors of this sort have a vital role to play in understanding speakers’ assessment of *wh-want* sentences. In particular, we believe that differences in the sensitivity of native and non-native speakers to the properties of individual lexical items lead to different expectations concerning the likelihood that the *wh* dependency will be resolved at *want* rather than at some later point in the sentence.

Native speakers

We hypothesize that the key external factor for NSs involves the relative frequency of the various *wh-want* patterns. As revealed by Ito’s (2005:121) study of the three most common uses of *want* in the Michigan Corpus of Academic Spoken English (MICASE), the <N TO> pattern is far more frequent than the <N N TO> pattern (and therefore presumably more likely to appear in *wh* questions). (12.14% of the uses of *want* in MICASE were not analyzable.)

(31)a. <N TO>: 63.99%

(e.g., They want to visit Mary.)

b. <N N TO>: 19.67%

(e.g., They want me to visit Mary.)

c. <N N>: 4.2%

(e.g., They want her.)

It is therefore quite possible that the NS processor, which has presumably become attuned to the tendencies manifested in the input, will initially bypass

want (or at least not strongly commit to a gap there) when processing *wh-want* patterns, placing its bets on a later opportunity to resolve the *wh* dependency.

Non-native speakers

We hypothesize that matters are very different for non-native speakers, who as a group will (we predict) routinely attempt to resolve the *wh* dependency in *wh-want* patterns at *want* rather than at the embedded verb.

- (32) The Early Resolution Hypothesis:
NNSs as a group exhibit a stronger expectation than do NSs that the *wh* dependency can be resolved at *want*.

As we will explain below, the propensity to resolve the *wh* dependency at *want* creates a situation that encourages over-acceptance of contraction patterns.

A simple way to test the Early Resolution Hypothesis makes use of the ‘filled gap’ paradigm (Stowe 1986) to investigate the following contrast.

- (33) *Wh* pattern in which the post-*want* position is occupied by a nominal:
(I wonder) who the teachers want *John* to help during the holidays.
- (34) *If* pattern in which the post-*want* position is occupied by a nominal:
(I wonder) if the teachers want *John* to help someone during the holidays.

The Early Resolution Hypothesis makes two predictions.

- i) Because native speakers do not anticipate a strong opportunity to resolve the *wh* dependency at *want*, they do not expect a gap in the post-verbal position. There will therefore be no difference in their reading profiles for (33) and (34), since the presence of an NP after *want* is equally (un)expected in either pattern.
- ii) Because second language learners seek an opportunity to resolve the *wh* dependency at *want*, they will exhibit a filled-gap effect in patterns such as (33). This is because their attempt to link the *wh* word to the second argument position in the verb’s grid will subsequently be contradicted by the presence of the verb’s true second argument (*John*), creating a filled-gap effect in the *wh*-pattern.

Materials

Twenty-four sets of contrasts such as the following were created.

(37) Sample test items

a. *who-want*:

I wonder who the mothers want the instructor to thank at the party.

b. *if-want*:

I wonder if the mothers want the instructor to thank John at the party.

The test items were distributed into four presentation lists using a Latin square design. Each participant read only one of the four lists. Seventy-two filler sentences were inserted into each list in pseudo-random order so that at least one filler sentence intervened between any two test sentences.⁴

Procedure

The experiment used a word-by-word self-paced reading task controlled by E-Prime software running on a PC computer. The reading task began with the presentation of a sentence masked with dots on the computer screen. By pressing one of the keys on the keyboard, the participants revealed the first word of the sentence. Pressing on the key again revealed the second word and turned the first word back into dots. The participants proceeded in this manner until reaching the end of the sentence, marked by a period. Upon reading the last word of the sentence, the participants pressed the key one final time, triggering the presentation of a simple query to test whether they had been attending to the task.⁵

The computer recorded the interval for each key press, thereby measuring the reading time of each word. Instructions and eight practice sentences preceded the experimental trials.

⁴ There were four experimental conditions in the original design—two involving *tell* and two involving *want*. Twenty-four of the fillers involved verbs with <N N to> argument structures and early resolution of the *wh* dependency (e.g., *I wonder who the commander of the US Army told/wanted t to look after Kevin*). The remaining 48 fillers involved verbs with <N to> argument (*hope, try, and want*), half involving *wh* dependencies and half consisting of *yes-no* questions.

⁵ The queries were in the form of fill-in-the-blank statements (with a choice of two answers) pertaining to an event involving the matrix verb, the embedded verb, or an adjunct phrase in the test sentence. The number of correct/incorrect responses and their relative position on the screen were counterbalanced.

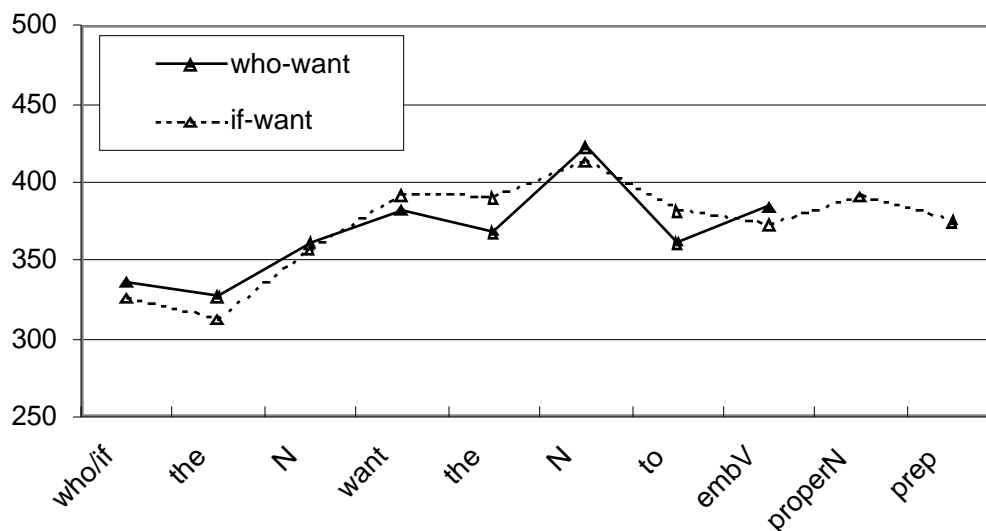
Data analysis

All data analyses were conducted separately for NSs and NNSs. The data from two native speakers and five non-native speakers were discarded either because of poor performance (below 70%) on the end-of-sentence comprehension task for test or filler items, or because they claimed to be NNSs but had received at least part of their primary or secondary education in English.

Analyses were conducted of the reading times of the remaining 36 native speakers and 36 non-native speakers for test items that had elicited correct responses in the comprehension task. Data points that fell more than 2.5 standard deviations from the region \times condition means for all participants were discarded. (This affected less than 2.3% of the NS data and less than 2.9% of the NNS data.) The means for each region \times condition were then calculated for each participant (subject analysis) and for each item (item analysis), and subjected to repeated measures ANOVAs.

6.3 Results

The results for the native speakers are summarized in Figure 1.



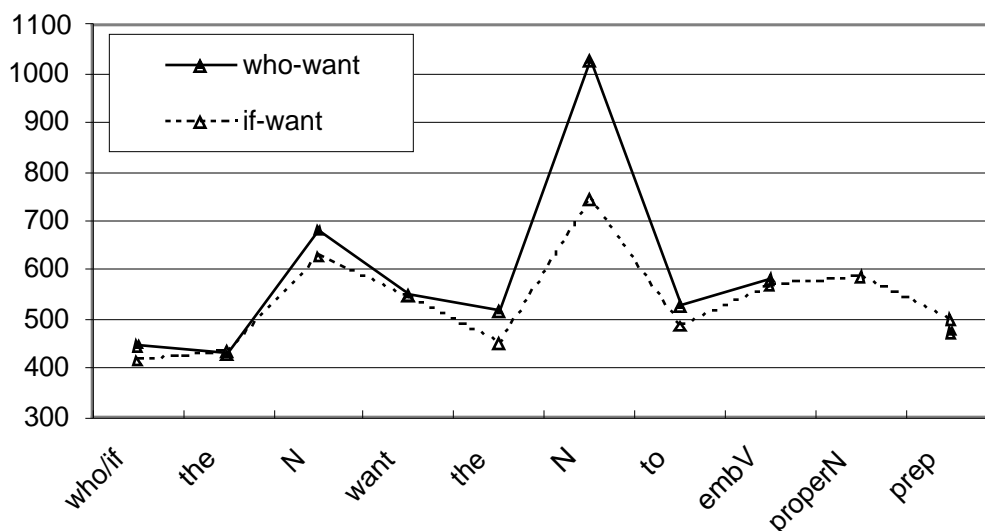
- a. (I wonder) who the mothers want the instructor to thank at the party.
 b. (I wonder) if the mothers want the instructor to thank John at the party.

Figure 1 Reading times (msec.) by native speakers

The key finding here is the absence of a significant increase ($F_s < 1$) in the reading time of the *who-want* pattern compared to the *if-want* pattern in region 6, the position occupied by the head of the post-*want* NP. There was, in other words, no filled gap effect. We interpret this as evidence that the NS processor is not

strongly anticipating an opportunity to link the *wh* word to an open position in the grid of *want*.

Figure 2 summarizes the performance of non-native speakers.



- a. (I wonder) who the mothers want the instructor to thank at the party.
 b. (I wonder) if the mothers want the instructor to thank John at the party.

Figure 2 Reading times (msec.) by non-native speakers

The key finding here is the significant increase in reading time in the post-*want* region in the *who-want* pattern compared to the *if-want* pattern (subject analysis: $F(1,35)=21.465$, $p < .001$; item analysis: $F(1,23)=13.43$, $p < .002$). This points toward a clear filled gap effect, suggesting that the L2 learners were anticipating an opportunity to resolve the *wh* dependency at *want*.⁶

6.4 Discussion

The finding that NNSs differ from NSs in attempting early resolution of the *wh* dependency in *wh-want* patterns is of considerable inherent interest, but the more important matter from our perspective has to do with whether it can shed light on our central research question, restated here.

- (38) Why do NNSs exhibit a greater propensity than NSs to be liberal in their acceptance of *want to* contraction?

⁶ A parallel set of contrasts involving *tell* (*I wonder who/if the mothers told the instructor to thank...*) revealed only a non-significant slowdown at *tell* in the *who* pattern.

NNSs who adopt this approach end up in the same category as ‘liberal’ native speakers, accepting sentences such as *Who do you wanna stay?*, even though phonological reduction takes place under non-optimal conditions.

There are currently no grounds for predicting which of these strategies individual second language learners will opt for—and indeed there is no apparent difference in the processing profiles of the L2 learners who differentiate between the two type of *want-to* patterns and those who do not. A judgment task administered right after the self-paced reading task yielded the following subgroupings of NNSs (using Ito’s criterion, as described in section 5; see table 1).

Table 3 Performance of NNSs on the grammaticality judgment test

Conservative	Overgeneralizers	Native-like
17	4	15

A two-way ANOVA reveals no difference for group, only the predicted difference for sentence type (the *who-want* pattern vs. the *if-want* pattern—the ‘filled gap’ effect).

Because of the small number of subjects who overgeneralized and because Ito’s criterion for acceptability (a mean rating of 3.2 out of 4) is, like all such criteria, somewhat arbitrary, we also tested for a correlation between the strength of the acceptability ratings and the reading times in region 5 and 6. Once again, no correlation was found.

These results are perfectly consistent with our hypothesis. The issue that we have been trying to address involves the question of why second language learners in general exhibit a greater propensity to accept *want-to* contraction in patterns such as (39) than do native speakers. A plausible answer to that question is now at hand. Because L2 learners routinely seek to resolve the *wh* dependency at *want*, (unlike native speakers) they find themselves in situations where they must choose between early resolution of the *wh* dependency and the naturalness of contraction, as we have just seen.

A preference for the naturalness of contraction leads to the rejection of (39), resulting in a native-like judgment of unacceptability. A preference for early resolution encourages the acceptance of (39)—the hallmark of ‘liberal’ speakers. And the avoidance of contraction altogether circumnavigates the dilemma, yielding the judgment profile characteristic of conservative speakers.

7. Concluding remarks

The phenomenon of *want-to* contraction is one of the most intriguing in all of language. The standard analysis is characterized by abstractness as well as elegance—it presupposes the existence of elements (traces) that are not only invisible, but must be distinguished from other invisible categories by an abstract Case feature. (As noted at the outset, neither PRO nor a non-Case-marked trace blocks contraction.) Not surprisingly, such accounts appeal to an innate principle of Universal Grammar to bridge the gap between the (trace-less and Case-less) input to which language learners are exposed and the exquisitely abstract grammatical constraint that the language appears to obey.

Matters are very different in the emergentist approach, which seeks to explain the properties of language by reference to the interaction of non-grammatical phenomena. In the case of the particular proposal we have made here, the key interaction involves two factors.

On the one hand, the computational system exhibits a propensity to resolve dependencies at the first opportunity, which we take to be motivated by the need to minimize the burden on working memory. On the other hand, there is the fact that contraction is most natural when the affected elements combine without delay (since the articulatory processes involved in phonetic fusion are, for physiological reasons, most common and most natural in rapid connected speech).

Matters are straightforward in patterns such as (41), where *want* combines with *to* without delay.

contraction is fully natural
↓
(41) Who_j do they_i want to see?
 <N_i TO>

But a conflict arises in (41), where the need to resolve the *wh* dependency at *want*, creates a delay that diminishes the naturalness of contraction.

contraction is less than natural for most native speakers
↓
(42) Who_j do they_i want to stay?
 <N_i N_j TO>

English native speakers typically reject contraction in this context, but a large number of NNSs accept it. Based on the experimental study reported in section 6, we have attributed this to a difference in expectations about where the *wh*

dependency is to be resolved in *want* questions. Whereas NSs place their bets on the embedded verb, NNSs expect an opportunity to resolve the *wh* dependency at *want* (the Early Resolution Hypothesis). A willingness to reconcile this expectation with the presence of contraction leads to liberal judgments concerning the acceptability of sentences such as *Who do they wanna stay?*

This idea in turn raises a new research question: why do second language learners come to have expectations about dependency resolution that differ from those of native speakers? A variety of possibilities suggest themselves, including limitations on working memory that favor early resolution of stored dependencies and a possible reduced sensitivity among second language learners to the distributional subtleties of the input.

In addition, it is worth noting that the vast majority of subjects in our study, as well as all the subjects in both the Kweon study and the Ito study, were native speakers of ‘*in situ* languages,’ in which *wh* words occur sentence-internally in argument positions. This raises the possibility that active filler effects might be stronger in speakers of such languages. Unfortunately, we are unable to pursue this matter here, as our study did not include native language type as a factor in the experimental design. Only 7 of our 36 subjects had a ‘movement-type’ native language—too small a number to permit statistically valid comparisons.

This is just one of the many issues that call out for attention as work continues on the emergentist thesis and its consequences for our understanding of second language acquisition.

References

- Bailes, Gregg. 2000. *Wanna* contraction in a British English environment. M.A. thesis. Department of Linguistics and English Language, University of Durham.
- Barss, Andrew. 1995. Extraction and contraction. *Linguistic Inquiry* 26, 681-94.
- Boeckx, Cedric. 2000. A note on contraction. *Linguistic Inquiry* 31, 357-66.
- Clahsen, Harald & Claudia Felser 2006a. Continuity and shallow structures in language processing. *Applied Psycholinguistics* 27, 107-126
- Clahsen, Harald & Claudia Felser 2006b. Grammatical processing in language learners. *Applied Psycholinguistics* 27, 3-42.
- Clifton, Charles & Frazier, Lyn. 1989. Comprehending sentences with long-distance dependencies. In G. N. Carlson & M. K. Tanenhaus (eds.). *Linguistic structure in language processing*. Dordrecht: Kluwer.

- Crain, Stephen & Janet Dean Fodor. 1985. How can grammars help processors? In D. Dowty, L. Karttunen, & A. Zwicky (eds.), *Natural language parsing: Psychological, computational and theoretical perspectives*. Cambridge University Press.
- Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar*. Cambridge, MA: MIT Press.
- De Vincenzi, Marica. 1991. *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.
- Elman, Jeffrey. 1999. The emergence of language: A conspiracy theory. In B. MacWhinney (ed.), *The emergence of language*, pp. 1–27, Mahwah, NJ: Erlbaum.
- Felser, Claudia & Leah Roberts. 2007. Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research* 23, 9-36.
- Fodor, Janet Dean. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9, 427-73.
- Garnsey, Susan, Michael Tanenhaus, & Robert Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistics Research* 18, 51–60.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, & Ivan Sag. 1985. *Generalized phrase structure grammar*. Oxford, UK: Blackwell.
- Hawkins, John. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75, 244–285.
- Hawkins, John. 2004. *Efficiency and complexity in grammar*. Oxford, UK: Oxford University Press.
- Ito, Yasuko. 2005. A psycholinguistic approach to *wanna* contraction in second language acquisition. Ph.D. dissertation, University of Hawaii at Manoa.
- Jaeggli, Osvaldo. 1980. Remarks on *to*-contraction. *Linguistic Inquiry* 11, 239-45.
- Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8, 63-100.
- Kweon, Soo-Ok. 2000. The acquisition of English contraction constraints by advanced Korean learners of English: Experimental studies on *wanna* contraction and auxiliary contraction. Ph.D. dissertation, University of Hawaii at Manoa.
- MacWhinney, Brian. 1999. Preface. In B. MacWhinney (ed.), *The emergence of language*, pp. ix–xvii, Mahwah, NJ: Erlbaum.
- Marinis, T., L. Roberts, C. Felser & H. Clahsen. 2005. Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27, 53-78.
- Menn, Lise. 2000. Babies, buzzsaws and blueprints: Commentary on review article by Sabbagh & Gelman. *Journal of Child Language* 27, 753–55.

- O'Grady, William. 1997. *Syntactic development*. Chicago: University of Chicago Press.
- O'Grady, William. 2001. An emergentist approach to syntax. Available at <http://www.ling.hawaii.edu/faculty/ogrady/>.
- O'Grady, William. 2005. *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Erlbaum.
- O'Grady, William. 2007. The emergentist thesis. This issue.
- Postal, Paul. 1974. *On raising: One rule of grammar and its theoretical implications*. Cambridge, MA: MIT Press.
- Postal, Paul & Geoffrey Pullum. 1982. The contraction debate. *Linguistic Inquiry* 13, 122-38.
- Pullum, Geoffrey. 1997. The morphological nature of English *to*-contraction. *Language* 73, 79-102.
- Steedman, Mark. 2000. *The syntactic process*. Cambridge, MA: MIT Press.
- Stowe, Laurie. 1986. Parsing *wh*-constructions. *Language and Cognitive Processes* 1, 227-246.
- Warren, Paul, Shari Speer, & Amy Schafer. 2003. *Wanna*-contraction and prosodic disambiguation in US and NZ English. *Wellington Working Papers in Linguistics* 15, 31-49.
- White, Lydia. 1989. *Universal Grammar and second language acquisition*. Philadelphia: John Benjamins.
- Williams, John, Peter Mobius, & Choonkyong Kim. 2001. Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics* 22, 509-40.