Learning to think like linguists: A think-aloud study of novice phonology students

Catherine Anderson

McMaster University

A key learning outcome for undergraduate linguistics courses is for students to learn to reason scientifically about language. This article presents the findings from a think-aloud study of undergraduates in an introductory linguistics course who were in the process of learning linguistic reasoning about phonology. I describe the students' developing concepts and make recommendations for instructors to help students develop fully formed linguistics concepts and the ability to think scientifically about language.*

Keywords: pedagogy, think-aloud, threshold concepts, novice thinking, disciplinary thinking, phonology

Every field has ways of thinking that characterize the discipline (Anderson & Hounsell 2007, Middendorf & Pace 2004, among many others), and the field of linguistics is no different. Kuiper (2011:183) asserts that '[w]hen you get a group of linguists together there is remarkable agreement about the domains that are central to linguistics'. In a survey of linguistics instructors in North America and Europe (Anderson 2012), I asked participants to rate a set of possible learning outcomes for an undergraduate linguistics degree. All seventy-one respondents agreed or strongly agreed that students should understand the principles of scientific reasoning as they apply to the study of linguistics. A core element of scientific reasoning in linguistics is the ability to think about language as observable data and to draw conclusions from one's empirical observations. In this article, I describe a study of beginning linguistics students' attempts to make observations and draw conclusions about the phonology of two unfamiliar languages. The findings provide insight for instructors into the nature of novice linguists' mental representations of key linguistic concepts such as phonemic contrast and allophonic variation.

Because of the pervasiveness of prescriptive attitudes about language, the ability to think empirically about language is no doubt a THRESHOLD CONCEPT for novice linguists. Meyer and Land's (2003) influential theory argues that threshold concepts, which are foundational to the ways of thinking within each discipline, are transformational for learners, 'opening up a new and previously inaccessible way of thinking about something' (2003:1). But this transformation is difficult and time consuming, and it can take weeks or months for novices to cross the threshold into a disciplinary way of thinking. During these weeks, Meyer and Land argue, students function in a liminal space, 'a suspended state in which understanding approximates to a kind of mimicry' (2003:10). The present study of students in an introductory course offers a view of novices' concepts within this liminal space, where students approach the threshold at different rates. From my findings, I offer some recommendations to instructors of introductory linguistics courses for supporting students in crossing the threshold to linguistic thinking.

1. The think-aloud method.

The THINK-ALOUD METHOD (Ericsson & Simon 1993, van Someren et al. 1994) offers researchers a useful tool for observing thinking

* I gratefully acknowledge support from the Faculty of Humanities at McMaster University and especially the work of two undergraduate research assistants, Laura Beaudin and Leslie Humphries, who collected and transcribed the data.

processes. Ordinarily, mental processes are hidden not only from an observer or researcher, but usually also from the thinker herself. The method involves simply asking participants to talk out loud about everything they think while they attempt to complete a task of some kind. The researcher audio-records the participants' utterances, transcribes them, and analyzes the resulting transcripts, which are referred to as PROTOCOLS within this research tradition. The think-aloud method is useful for studying learning for several reasons. Blackwell and colleagues (1985) give evidence that a think-aloud exercise generates significantly more data than asking participants to list their thoughts retrospectively, and Genest and Turk (1981) note that statements generated via the think-aloud method are relatively unlikely to be subject to participants' causal inferences about their own thoughts, unlike a retrospective listing of thoughts. The protocols resulting from a think-aloud reveal not just the mental processes involved in the task, but also the order in which those processes unfold over time (Whitney & Budd 1996).

Researchers have used the think-aloud method to investigate general problem-solving (Newell & Simon 1972) and inferencing processes while reading texts (Trabasso & Magliano 1996, Whitney & Budd 1996). Although the level of self-consciousness required to describe one's thinking aloud can disrupt cognitive processes for tasks that are overlearned or automatized (Genest & Turk 1981), a think-aloud is ideal for observing the mental processes of beginners, since the very process of learning a new task involves conscious reflection on how to carry out the task (Chi et al. 1994, Ericsson et al. 1993). For that reason, researchers of teaching and learning have employed the method to investigate how students begin to develop disciplinary thinking in history (Wineburg 1991), political science (Bernstein 2010), and mechanical engineering (Miller-Young 2013). The method has been particularly useful for observing the development of clinical reasoning across many fields in the health professions (Banning 2008, Forsberg et al. 2014, Kuipers & Kassirer 1984, Wainwright & McGinnis 2009). The present article extends the existing SCHOLARSHIP OF TEACHING AND LEARNING (SoTL) literature to consider the cognitive processes and conceptual representations of novice students of linguistics.

1.1. PARTICIPANTS. A small number of participants is usually considered sufficient for a think-aloud study, on the assumption that there are only a few qualitatively different ways that students will understand (or misunderstand) course concepts (Marton 1986). Seven McMaster University undergraduate students (ages seventeen to twenty-five) participated in this study.[1] Six were novices enrolled in the first semester of a two-semester Introduction to Linguistics sequence[2] and had not taken any other linguistics classes. The seventh student was a fourth-year linguistics major who had completed all of the core theoretical classes in the program and was working as a teaching assistant (TA) for the Introduction course.

The participants were recruited via the department's experiment participation software. As compensation they earned one hour of participation credit toward a linguistics course. Two research assistants (RAs) handled recruiting and compensation and conducted the sessions, so that participants' identities were unknown to me in my dual roles of researcher and course instructor. All participants gave informed consent to participate.

[1] We did not collect gender information from our participants. I use feminine pronouns to refer to students throughout this article because the large majority of linguistics students at McMaster are women.

[2] The first-semester course deals with phonetics, phonology, and morphology, while the second semester includes syntax, semantics, and psycholinguistics. The language of instruction at McMaster is English.

1.2. MATERIALS. Each participant attempted two exercises, reproduced in Tables 1 and 2.

Here is a set of data transcribed from Lebanese Arabic.

[btækli]	'you eat'	[ʔɪntu]	'you' (*plural*)
[kʊɾsi]	'chair'	[bɪnt]	'girl'
[ʔɪnti]	'you' (*singular*)	[mɪtɪl]	'like'
[maʕi]	'with me'	[ʔɪli]	'to me'
[fi]	'there is'	[bɪʔi]	'he stayed'

Consider the vowels [i] and [ɪ]. Are they two separate phonemes, or are they allophones of one phoneme? What is the evidence for your conclusions?

TABLE 1. Lebanese Arabic exercise, from Cowan & Rakušan 1987.

Here is a set of data transcribed from Angas, a language spoken in Nigeria. (*Remember that this diacritic ̥ indicates that a sound is voiceless. The diacritic ʷ indicates that a consonant is rounded. The symbol* [ɓ] *stands for a voiced implosive bilabial stop.*)

[mut]	'to die'	[pampam̥]	'bread'
[ŋgak]	'snake'	[ntaŋzum̥]	'wasp'
[ndarm̥]	'bark'	[nfʷarm̥]	'head cold'
[nuŋ̊]	'to ripen'	[mɓɛlm̥]	'to lick'
[mbaŋga]	'drum'	[tam̥]	'bench'
[dondoŋ̊]	'yesterday'	[potiŋ̊]	'sky'
[dɛŋ̊]	'to drag'		

Consider the nasal consonants in this data set. How many separate phonemes are there? How many of them are allophones? What is the evidence for your conclusions?

TABLE 2. Angas exercise, adapted from Halle & Clements 1983.

Five students attempted the Angas exercise followed by the Lebanese Arabic exercise. One novice student attempted Angas first, moved on to attempt Lebanese Arabic, then returned to the Angas exercise. One novice student, the first to participate, attempted an Inuktitut exercise instead of Angas. Since the course's TAs then used that same Inuktitut exercise in class, we replaced it with Angas for all subsequent participants, and excluded the data from the Inuktitut exercise from our analysis.

1.3. PROCEDURE. The think-aloud sessions took place during weeks 8 to 12 of the thirteen-week semester. The topics of phonemic contrast and allophonic variation had been introduced in week 7 of the semester.

Two RAs conducted all sessions in a quiet room on campus. At the beginning of each session, participants read the letter of information, signed the consent form, and filled out a brief participant information form.

Each participant wore a head-mounted microphone connected to a computer running audio-recording software. The RAs invited the participant to think aloud by saying the following.

(1) Please complete the exercise while describing your thinking out loud. Try to explain what you're thinking while you work. We will give you prompts if you stop talking while you are working, but we cannot help you if you get stuck. (R)[3]

[3] Excerpts from the protocols are labeled with a code indicating the speaker: R for a Research assistant, N for a Novice student, and E for the Experienced student.

The RAs asked each participant to read the exercise instructions aloud while they checked the recording levels. After reading the instructions, participants began to describe their thinking. They also wrote on the exercise paper as if they were completing a graded assignment. If participants stopped speaking, the RAs prompted them with a question or a reminder to keep thinking out loud. Most prompts were simple reminders to keep talking, for example, 'What are you thinking about now?'. Some of the prompts by the RAs included coaching such as a TA might provide, as in 2.

(2) Going back to the nasal consonants, have you determined how many separate phonemes there are? (R)

Charters (2003:73) has argued that, '[i]deally, participants in a think-aloud study should not need any coaching but should enunciate their inner speech spontaneously'. It is possible that the RAs' coaching prompts made the protocol unrealistic as a representation of the students' abilities. In other words, if the student were attempting to complete this problem unassisted on an exam, she might not get any further without such a prompt. But my present interest is not just in whether students can reach the right answer unassisted, but also in what their mental processes are while they are learning how to solve a problem. The RAs' occasional coaching prompts gave us a fuller data set to illustrate the thinking that would occur in a real-life learning situation, where students often receive coaching from peers, TAs, and instructors.

If a participant was unable to reach a solution, the RAs concluded the session by asking some retrospective questions about the exercise, such as 3 and 4.

(3) Was there anything that you struggled with? (R)

(4) What do you think would be helpful for you to complete this problem? (R)

The session durations ranged from fourteen to thirty-seven minutes (average twenty-four minutes). After each session the RAs transcribed the participant's utterances into text files.

2. Protocols and analysis. The RAs transcribed seven sessions of the Lebanese Arabic exercise (six novice, one expert) and six sessions of the Angas exercise (five novice, one expert). We excluded the Inuktitut exercise from the present analysis in the interest of greater comparability across protocols. This resulted in a total of thirteen protocols (eleven novice, two expert) for analysis and coding.

In my coding of the transcripts, I used an inductive approach that did not begin with a hypothesis or preconceived set of codes; rather, I allowed the code system to emerge from my observations of the protocols. Nevertheless, my approach was not theory-neutral because my observations and inferences were guided by a large literature on threshold concepts and disciplinary thinking (Meyer & Land 2003, 2005, 2006, Middendorf & Pace 2004, Wismath et al. 2015), by the cognitive psychology literature on problem solving (Ayres 1993, Catrambone 1998, Chi et al. 1989, Ericsson 2003, Getzels 1982, Larkin et al. 1980, Lesgold 1988, Newell & Simon 1972, Owen & Sweller 1985, Pretz et al. 2003, Reed 1987, Reed & Bolstad 1991), and of course by my own experience of more than a decade of teaching introductory linguistics.

2.1. Solving a phonology problem. The typical data set from an unfamiliar language, accompanied by instructions to consider a pair of phonetic segments, is generally referred to by linguists as a *problem set*. Although in my classes I use the term *exercise*, these exercises are examples of classic problems: they have clear goals and a well-defined path to solution. Exercises of this kind are, in Getzels's (1982) terms, the type of problems that are presented to students (as opposed to *discovered* or *created* by the problem-solvers).

In a useful review, Pretz and colleagues (2003) summarize a large literature on problem solving by listing the common stages involved in solving a problem.

(i) Recognize or identify the problem.
(ii) Define and represent the problem mentally.
(iii) Develop a solution strategy.
(iv) Organize [one's] knowledge about the problem.
(v) Allocate mental and physical resources for solving the problem.
(vi) Monitor … progress toward the goal.
(vii) Evaluate the solution for accuracy. (Pretz et al. 2003:3–4)

They also note that these stages are not necessarily sequential, nor are all stages present in every problem-solving situation. Since the phonology problems under consideration are presented to students, stage (i) of the cycle is not relevant in the present situation. Likewise, because the novice students are total beginners at doing linguistics, most are not yet independently capable of developing a solution strategy (stage (iii)). Instead, most attempt to follow an algorithm or 'recipe' learned in class. In this introductory class, phonology exercises were modeled with the strategy shown in 5.

(5) a. Look first for minimal pairs.
b. If you find minimal pairs, conclude that the segments are separate phonemes.
c. If you don't find minimal pairs, consider the distribution of the segments.
d. If you find complementary distribution, conclude that the segments are allophones of one phoneme.

As we shall see from the protocol data, many, but not all, students were able to follow this recipe, but did not always reach a correct solution. Some errors resulted from applying the strategy incorrectly (stage (iii)), while many arose earlier, from students' mental representation of the problem at stage (ii). In some cases, students executed the steps according to the recipe but still failed to reach a correct solution because of gaps or errors in the organization of their knowledge about the problem (stage (iv)).

2.2. The 'correct' solutions. The Lebanese Arabic data set was the simpler of the two, because it questioned only a single pair of segments. The evidence confirms that the students found this exercise easier: four of the six novices reached a correct solution, as did the experienced student, and the average time to complete this exercise for the novices was 7;35 minutes (as compared to 12;50 minutes for the Angas exercise).

For an expert linguist, a glance at the data is sufficient to observe that [i] appears word-finally and [ɪ] appears elsewhere, and to conclude therefore that [i] and [ɪ] are allophones of one phoneme. Indeed, the experienced undergraduate participant reached this conclusion almost immediately. Item 6 shows her first utterance after reading the instructions.

(6) I'm observing the data to check what positions the [i] and the [ɪ] are falling into. Initial observations are that [i] are happening only word final and that [ɪ] is happening medial, medial, medial, medial, medial. Well, they're never appearing in the same environments. So, based on like a really basic basic observation, you could already state that they're in complementary distribution. (E53)

After three more sentences during which she conducted a more detailed examination of the phonetic environment of each segment, she concluded as follows.

(7) It kinda looks like they are indeed allophones of one phoneme because they're never appearing in the same environments. (E53)

For a novice linguist, however, reaching that conclusion would require several steps. To reach a successful solution, that is, one that would earn full points on an exam, a student would need to do the following.

(8) a. Activate the solution strategy described above in 5 from memory;
b. locate the relevant segments in the data set;
c. compare word forms to discover whether the segments appear in minimal pairs (and in this exercise, avoid being distracted by an irrelevant minimal pair);
d. access phonetic knowledge from memory to describe the environments in which each segment occurs (optionally using an environment chart);
e. recognize the distribution of the segments as complementary;
f. use that recognition to conclude that the two segments are allophones of one phoneme.

The Angas exercise does not include much more data than the Lebanese Arabic data set (thirteen words instead of ten), but it is conceptually more complex because it asks students to consider not just a single pair of segments, but the natural class of nasal consonants. The steps to solve the problem are essentially the same as the steps described above in 8 for Lebanese Arabic, but arriving at the correct solution requires the additional conceptual leap of recognizing that the entire natural class participates in the allophonic variation, such that the voiceless nasals [m̥ n̥ ŋ̊] appear word-finally and their voiced counterparts [m n ŋ] appear elsewhere.

Within the natural class, there are no minimal pairs in the data set to provide conclusive evidence that /m/, /n/, and /ŋ/ are three separate phonemes, but there is one near-minimal pair ([mut] vs. [nuŋ̊]) for the /m/-/n/ contrast. However, this small data set also includes some evidence that could be interpreted as showing a single underlying nasal changing its place of articulation according to the following consonant, since in many of the words provided, [m] precedes bilabial segments, [n] precedes alveolars, and the single instance of word-initial [ŋ] precedes [g]. In short, the evidence for phonemic contrasts between [m], [n], and [ŋ] is subtle and would likely therefore be quite challenging for a novice.

2.3. Students' approaches to the problems. Most of the steps sketched above in 8 are themselves conceptually quite complex: they rely not just on recognition and recall but on rich mental representations of core linguistic concepts such as phonetic features, natural classes, minimal pairs, and allophonic variation. Chi and her collaborators (1989) showed that students' abilities to solve problems arise from robust, complete knowledge representations. Indeed, the chief goal of most introductory courses in any discipline is for students to develop knowledge representations characteristic of that discipline (Anderson & Hounsell 2007, Middendorf & Pace 2004, Pace & Middendorf 2004). Since our study targeted students at the beginning of their linguistics studies, we expected to find that many novice students had fragmentary representations of linguistic concepts. The protocols from two novice students (N84 and N53) who were unable to reach a correct solution for either exercise indicate that these students' knowledge representations were confused and lacking detail. Perhaps more interesting to instructors, the protocols from successful solutions (even from the experienced student) reveal some misconceptions that persisted in students' concepts even when they were able to reach technically correct solutions.

Representing the problem. For both exercises, the protocols indicate that students spent little time in forming a mental representation of the problem (stage (ii) from Pretz et al. 2003), and instead very quickly began employing the algorithm learned in class. Each of the following quotes is a student's first utterance after reading the instructions.

(9) So looking at the nasal consonants, how many separate phonemes of the nasal consonants are there? I'm going to look at the environments that they are in, so I'm going to make—So I'm going to put the nasal consonants in a chart and then I'm going to put the environment of each. (N51)

(10) Well I'll probably draw a chart, like how I started, but what do you mean by the nasal consonants? (N39)

Evidence from a variety of domains shows that expert problem-solvers devote proportionally greater amounts of time than novices to setting up a mental representation of a problem, and proportionally less time executing the solution (Chase & Simon 1973, Larkin et al. 1980, Lesgold et al. 1988). The protocol statements in 9 and 10 are examples of typical novice behavior: the students attempted a strategy without creating a clear representation of the problem. In fact, item 10 indicates that student N39 had not marshalled enough phonetics knowledge to know which segments to consider but had already begun to create an environment chart.

By contrast, students with a clear representation of the problem were able to apply the problem-solving strategy efficiently. In 11, for example, student N76 was clear that the focus of the problem was nasal consonants.

(11) N: So if there is no minimal pair then we go to figure out the environment the syllable is in.
R: How would you do that?
N: Create a chart. So there are some words without [m], so I will probably just ignore them. (N76)

This protocol statement suggests further that she had adopted a strategy of focusing on one nasal segment at a time, with the result that she was not distracted by irrelevant data. In contrast, without a clear representation of the structure of the problem, student N53 did not know what data to attend to.

(12) N: So I might read the words again and just list down on the paper where the separate phonemes are. I guess I would start with the first word [mut], so there would be three here.
R: Are any of those nasal phonemes?
N: No, are we just looking for nasal phonemes? (N53)

Excerpt 12 suggests that the student was attempting to execute step (c) from the recipe in 5, 'consider the distribution', but had not formed a representation of which segments were under consideration, so spent time on irrelevant details.

From an expert linguist's point of view, the environment chart is a tool for organizing one's knowledge about the problem. An environment chart can make it easier to observe patterns in the distribution of a particular segment. But the protocols suggest that many students perceived the environment chart not just as a useful step in the solution strategy, but as an end in itself. A vital challenge for instructors, therefore, is to make clear the difference between tools for observing data on the one hand, and ways of drawing conclusions about data on the other.

In contrast to the novices who jumped into the solution strategy right away, the experienced student invested time to represent the problem space mentally, as example 13 illustrates.

(13) E: Okay, so consonants involve [m], [ŋ], [n]. Mmm, okay. And voiceless [m̥], voiceless [n̥] and voiceless [ŋ̊]. (*whispering, reading instructions*)

How many separate phonemes are there? How many of them are allophones? What is your evidence for your conclusions? So let us put together our common ones.

R: What are you doing right now?

E: I am reading through all of the words and categorizing the nasals that match, that, at least in their surface forms, look similar to one another so I can see if there are certain environments that particular nasals are popping up in. (E53)

Student E53 first accessed her declarative knowledge about phonetics to determine which segments to consider, then started to identify the distribution of each. Her statement 'let us put together our common ones' suggests that she has a mental representation of a phonetic segment as a linguistic object with a predictable behavior. In fact, this student's conceptual representations of the notions of ALLOPHONIC VARIATION and DISTRIBUTION were robust enough that she not only had no need to make an environment chart to observe the relevant distributions, but also could not even remember what an environment chart is when the RA prompted her at the end of the think-aloud session.

These protocol statements from the outset of the problem-solving sessions show hints of students' incipient linguistic concepts. They also show that beginner students were encountering a METACOGNITIVE threshold concept. Wismath and colleagues concluded from their study of undergraduates' metacognitive problem-solving abilities that the third of three threshold concepts in problem solving 'involves the realization of the importance of careful and complete modeling of a problem before one plunges into tackling it' (Wismath et al. 2015:69). Since students in that study showed evidence of reaching this threshold only toward the end of a course that concentrated entirely on problem solving, it is not surprising that first-year linguistics students had not yet achieved it. Instructors who offer a recipe for solving phonology problems might profitably include an early step in the recipe such as 'think about what you already know' to remind students to invest time in representing the problem before starting to solve it.

LOOKING FOR MINIMAL PAIRS. The algorithm (5) that students practiced in class begins with a search for minimal pairs. Looking for minimal pairs is of course not logically necessary in solving a phonology problem. I advise beginner students to start with this step because when they are scanning a data set, it can be easier to locate a minimal pair than to notice patterns in the distributions of segments. About half of the novice protocols (six of eleven) indicate that the student started the problem by searching for minimal pairs, such as in 14.

(14) First try to find the minimal pairs. I don't think there are any minimal—there is one near minimal—not really. Try to find the [i] and [ɪ]. So circle all of the [i] in the words. … I'm just trying to find and label them all first. I don't think there is minimal pairs. So I will just start with trying to find the environments they are in. (N76)

In the seven protocols that did not mention searching for minimal pairs, it is possible that participants either forgot this step or were confused about it, like student N78.

(15) I'm going to start with similar words and see if they contrast. So, similar words, similar words, come on. Usually when I look for these words, I usually only look for the first letter and then I go from there. They showed me in class to make a table, so I'm making a table. On one side there's [ɪ], the stressed [i] and [ɪ]. Then I'll put the words that are similar beside each other. (N78)

This portion of student N78's transcript indicates some confusion among the steps in the recipe. Although she knew that searching for minimal pairs and making an environment chart were both steps in the recipe, her mental representation was not clear on the functions of the two different kinds of evidence.

But for many of the students who appeared to skip the search for minimal pairs, it is plausible that they could have actually performed a rapid scan and, not finding any minimal pairs, moved on immediately to observing phonetic environments without even commenting aloud on the minimal-pair search. In other words, not mentioning the minimal-pair step of the recipe out loud does not necessarily indicate that they skipped that step.

As mentioned above, the evidence within the Angas data set for phonemic contrasts between [m], [n], and [ŋ] is sparse. One novice student observed the absence of minimal pairs.

(16) Okay, so three nasals are separate phonemes. But there are no minimal pairs. I'll need minimal pairs to support it. But they are in the same environment. (N76)

This student had a clear representation of what kind of evidence would support a conclusion of separate phonemes: she would have preferred the strong evidence of minimal pairs but was willing to draw a conclusion based on the sounds' similar distributions.

But many students, even the experienced student, simply assumed that [m], [n], and [ŋ] are three separate phonemes without considering the evidence.

(17) In this case, there would be three separate phonemes. And um, how many of them are allophones? We've got, okay, there are six different segments that are appearing, three of which are voiced, three of which are voiceless. And I'm pretty sure that you've got at least one separate allophone per phoneme, so the phoneme for example would be [m], [n], and [ŋ], and then their allophones underneath them. (E53)

Even though the assumption appears to be correct in the case of this data set, making the assumption in the absence of evidence is an error in linguistic thinking. No doubt for some students this assumption arises from interference from their native language English, where [m], [n], and [ŋ] are indeed three separate phonemes. Students in the introductory course often encounter such interference when they are exposed to data from languages whose phonology differs from English. For example, voicing allophony in stops in Swampy Cree is often particularly challenging for students to recognize because the allophones are [p]-[b], [t]-[d], and [k]-[g], each pair of which is contrastive in English.

In addition to interference from English, the protocol statements also reveal other flaws in students' conceptual representations of phonemic contrast. The IPA notation itself led some students to make assumptions about phonemes and allophones. Example 18 suggests that student N76 was under the impression that allophones are indicated by diacritics.

(18) [ŋ] happens here, here, and here. But they don't have allophones. How many of them are allophones, but they don't have that circle there.[4] There is no characteristic to distinguish those. So do I still need to consider them? (N76)

By that kind of reasoning, segments represented by separate IPA symbols would automatically be assumed to be separate phonemes.

[4] The diacritic for voicelessness.

Student N84's protocol statements reveal the weaknesses in her concept of phonemic contrast. One source of confusion was between phonetic difference and phonemic contrast.

(19) So then in this case, um, [n] is only found at the beginning but [m] is also found at the beginning so that doesn't really distinguish them. So I'd say that they're two separate phonemes because they're also distinguished by the fact that [n] is alveolar and [m] is bilabial so they're different in more than two—in more than one criteria. (N84)

The confusion about the idea of contrast emerged from uncertainty about the nature of minimal pairs. When Student N84 was struggling with the problem, the RAs offered some coaching prompts.

(20) R: OK, so what's a minimal pair?
N: If they're only similar in one criteria?
R: What's a criteria in this case?
N: I don't know if has to be that they're similar in more than one way or if they're different in more than one way, then they're considered a minimal pair. I know it has something to do … something along those lines.
R: OK, can you give me an example of a minimal pair?
N: So there was like *stun* and *spun*.[5] And I believe that was considered a minimal pair. *stun* and *spun*—because—I'm checking to see if they're like voiced or voiceless—[p], [t]—so they're both voiceless. So they're only different in the manner of articulation, I suppose, so like [p] is bilabial and [t] is alveolar, so that's the only way, like, they differ, so then, that—that's why they're minimal pair. (N84)

N84's responses reveal that her concept of a minimal pair was a pair of segments that differ minimally. This mistaken impression might arise from parallels between the concepts: minimal pairs are pairs of words that differ in only one segment, and phonology problems often focus on pairs of segments that differ in only one feature. If a student's understanding of minimal pairs is flawed in this way, then a consequent confusion between phonemes and allophones is almost inevitable.

Identifying distributions. The algorithm in 5 then directs students to consider the distribution of the relevant segments. For the Lebanese Arabic problem, all of the students, novice and experienced, arrived at a correct or nearly correct description of the distribution of the segments. All recognized that [i] appeared only word-finally. The experienced student and five of the six novices recognized the distribution of [ɪ] as elsewhere relative to the word-final distribution of [i], though not always using the term *elsewhere*. For example, participant N39 clearly understood that the underlying form has the more general distribution, even though she did not use the term *elsewhere*.

(21) Because the vowel, the tense vowel, appears to occur at the end of the word always and it doesn't show any other places where it could occur. And the lax vowel, there isn't a situation where it occurs at the end of the word, so I'm wondering if they pronounce it as the—they pronounce the lax vowel usually—I'm wondering if it's usually the lax vowel and at the end of the word it just turns into a tense vowel. (N39)

[5] From a recent quiz.

Student N84, who did not recognize the elsewhere distribution of [ɪ], showed a symptom common to many novices, describing the environment in too much phonetic detail and thereby missing the relevant pattern.

(22) Yeah like for [ɪ] you'd only find it after a consonant. You wouldn't really find it like, elsewhere, like it's only found in the middle of a word, like you don't see it at the beginning of a word, like you only find it in that one, like in between two consonants, and I'm guessing after a stop. (N84)

Recognizing meaningful patterns is an attribute of expertise (Chase & Simon 1973, Lesgold et al. 1988). Novice students' abilities to recognize distribution patterns will likely grow stronger as they gain more exposure to examples of allophonic variation across a variety of languages.

Identifying the distribution of allophones in the Angas data set caused students somewhat greater difficulty, no doubt because of the additional challenge of considering three pairs of nasal consonants rather than a single pair of segments. One novice student did not arrive at an accurate description of the distribution of the nasals. Two novices accurately described the pattern of complementary distribution for each of the three pairs of segments. The experienced student and only two of the novices described the distribution as applying to the natural class of nasal consonants, with voiceless nasals appearing word-finally and voiced nasals elsewhere.

These results suggest that the possibility of allophonic variation occurring within a natural class (and not just within a single pair of segments) is a conceptual move that takes some time for students to achieve. I predicted above that, as students gain exposure to multiple exemplars, their abilities to recognize conditioning environments would grow stronger. The same is likely for their abilities to recognize natural classes. The evidence from the protocols suggests that some students in this study were on the verge of taking that conceptual step but at the same time exercised a degree of caution about extending their conclusions too broadly. In 23, participant N39 had described the distribution of [m] and then turned her attention to [n].

(23) For the voiceless [n̥] I'm not really sure because there's only one so I'm a little confused—(*whispers*)—Um, I think I'll conclude that they're allophones because they don't occur in the same environment. And the voiceless [n̥] only appears to occur at the end of a word because the voiced [n]s don't occur at the end of the word, so I think I'll just end there. (N39)

This statement illustrates the tension in the novice conceptual representation: on the one hand, the student was reluctant to draw an inference about [n] and [n̥] from only a single data point for voiceless [n̥]. But on the other hand, she had noticed that the distribution of [n] and [n̥] is consistent with that of the other members of the natural class. This particular student was convinced by the evidence from [m] and [ŋ] that [n] also participates in the allophonic variation. But students with more fragile representations for the concept of a natural class were constrained by the paucity of data for [n] itself and were only willing to draw conclusions about [m] and [ŋ] individually.

Drawing conclusions. As we saw above, most of the students gave correct or nearly correct descriptions of the distribution of the allophones [i]-[ɪ] in Lebanese Arabic and of the voiced and voiceless nasals in Angas. In other words, they were largely successful in following the observation steps of the recipe. Let us now consider how they interpreted their observations.

For Lebanese Arabic, two novice students reached the wrong conclusion.

(24) I'd assume that they're separate phonemes then, because you wouldn't find one in one case and you wouldn't find the other in the other case. (N84)

(25) So the small [i] and the capital [ɪ] are two different phonemes, because the small [i] only appears at the end, and the capital [ɪ] never appears in the same environment as the small [i]. (N53)

It is possible that these incorrect solutions arise from simple terminology confusion. Instructors know that it often takes some time for beginners to remember the difference between the unfamiliar terms *phoneme* and *allophone*. But for students N84 and N53, the confusion in vocabulary appeared to be connected to a deeper confusion in the concepts. These same two students also struggled with the Angas problem: N53 was unable to reach any solution for Angas, while N84 reached a partial solution. And both expressed hesitation about their conclusions.

(26) R: What do you think would be helpful for you to complete this problem?
N: Um, for sure my notes to see, like again, to tell if I'm right or wrong about what's a phoneme, to assure myself of the definition of a phoneme. (N84)

(27) I think this is the best conclusion I can make at the moment, so I'll probably have to review my notes before the final or something. (N53)

In fact, even students who solved the problem correctly sometimes showed lingering confusion in the conceptual representation, as this example from student N78 indicates.

(28) Yep, allophones. I just have to show evidence though. So yeah, I think they're allophones of one phoneme. Why? I can't exactly tell you why. Can I give you evidence for why they're not phonemes? Does that work? … They're not phonemes, because they have no minimal pairs. Basically that's all I could say. They're distributed differently. Pretty sure that means—yeah I think, I don't remember completely. I'm writing that they're allophones of one phoneme, evidence of my conclusion, they're not phonemes because they have no minimal pairs. (N78)

Example 28 suggests that the student followed the algorithm correctly and indeed reached a solution that would probably earn full or nearly full points on a test. Yet her concept of allophony did not seem to be connected to the idea of phonetically conditioned variation: she just knew that if there are no minimal pairs then the segments are probably allophones. This example echoes findings from freshman chemistry, where students scored higher on algorithmic questions that used formulae than on conceptual questions on the same topics (Cracolice et al. 2008, Nurrenbern & Pickering 1987). This particular example reminds me as instructor to design assessment questions on tests that measure conceptual understanding without relying exclusively on students' abilities to execute an algorithm.

Student N51 gave a textbook correct solution to the Lebanese Arabic problem, identifying complementary distribution easily and concluding that the two segments were allophones of a single phoneme. But the RA's final question then revealed a fundamental misunderstanding.

(29) R: Did you have any difficulties with this one?
N: The fact that there were no minimal pairs, it was difficult to conclude exactly that they are two separate phonemes. (N51)

Although she had just concluded that [i] and [ɪ] were allophones of one phoneme, her final comment suggests that Student N51 expected to find evidence for both conclusions within the data set, and possibly even for the same pair of segments. I propose that this is another conceptual leap (maybe even a threshold concept): for a truly robust representation of the concept of allophony, students must understand that segments that vary allophonically are unlikely to contrast with each other in the data for a given language.[6]

3. RECOMMENDATIONS FOR INSTRUCTORS. This study has shown that the think-aloud method is useful for gaining insight into students' conceptual representations while they traverse the liminal space before crossing the threshold to linguistic thinking. Not only is this method useful for SoTL researchers, but it can also be a valuable tool for instructors and TAs to support students in one-on-one settings. Ann Bunger, another instructor of introductory linguistics, reported to me that she often uses a think-aloud exercise with students during her office hours to identify students' points of confusion (p.c., October 2015).

3.1. LEARNING TO THINK LIKE LINGUISTS. The protocols from the present study have revealed several areas where beginning learners of phonology have conceptual representations that differ from those of experienced linguists. Although this empirical study was small, its findings ring true with my anecdotal observations over many years of teaching this course. Bernstein (2010:54) argues, 'if experts do something that novices do not … , then instructors should strongly consider sharing, and scaffolding, these behaviors in class'. Let us therefore consider what expert-like habits an instructor can scaffold in an introductory class.

SET UP THE PROBLEM. We know from domains as diverse as physics, chess, and radiology that, when faced with a problem, experts spend a sizeable proportion of their time setting up a mental representation of the problem before initiating a solution strategy (Chase & Simon 1973, Larkin et al. 1980, Lesgold et al. 1988). The protocols showed that novices spent relatively little time on this step. In my course, I have modified the recipe in 5 to include a reminder about the importance of this step: 'Look at the segments you're comparing. What are their articulatory descriptions? What is the difference between them? What natural class(es) do they belong to?'

DISTINGUISH BETWEEN OBSERVATIONS AND CONCLUSIONS. Many students in this study seemed to perceive the environment chart listing the distributions of the segments as the solution to the problem. Since a fundamental goal of the course is to develop students' scientific reasoning about language, it is vital for students to begin to grasp the distinction between empirical observations about language and the conclusions linguists draw from those observations. This distinction can be especially challenging in the field of linguistics, where so many students begin the course with firmly held prescriptive beliefs about language.

A good deal of pedagogical research has explored the use of analogy in learning abstract scientific concepts (Clement 1998, Gentner & Gentner 1983, Gigerenzer & Hoffrage 1995, Vishton 2005). To support students in learning to reason scientifically about language, I often draw parallels between linguistics and fields that students more readily recognize as 'science'. For example, ornithologists OBSERVED a decline in the population of the Northern Spotted Owl and CONCLUDED that climate change and habitat loss threaten the species (Dugger et al. 2015). Likewise, sociolinguists OBSERVED a

[6] A referee points out that pairs of segments like [n] and [m], which are otherwise contrastive, may be allophones in a neutralizing environment, such as the English morpheme *un-*, which is [ʌn] in *undress* but [ʌm] in *unbelievable*. This subtle level of variation is beyond the scope of my introductory course. To reinforce the conceptual difference between allophonic variation and phonemic contrast for the learner, I avoid including such data in the exercises for this course.

rapid increase in the use of quotative *be like* among Canadians under thirty and CONCLUDED that not only adolescents but also adults participate in ongoing language change (Tagliamonte & D'Arcy 2007). In addition to these analogies, when the class engages in problem-solving activities I highlight the difference between the observations—minimal pairs and distributions in phonology, grammaticality judgments in syntax—and the conclusions that these observations allow us to draw about the mental grammar. For phonology specifically, I have reframed the algorithm in 5 to include the questions 'What do you observe about the segments in the data set?' and 'What can you conclude from your observations?'.

DEVELOP ROBUST REPRESENTATIONS. Just as language learners develop robust phoneme categories by exposure to multiple exemplars (Lively et al. 1993, Maye et al. 2002, Werker & Tees 1984), learners in many domains develop conceptual representations from experience with a wide variety of examples. (See, for example, the chapters in Chi et al. 1988.) The protocols make it clear that many beginning linguists have not yet developed robust representations for concepts like natural classes and conditioning environments. In other words, they have not yet seen enough examples to have linguist-like intuitions about what kind of allophonic variation is likely and what environments will likely lead to such variation. No doubt most instructors of introductory courses already incorporate many practice exercises into class sessions and assignments. The present findings confirm that it is vital for students to encounter multiple examples of phonological phenomena within actual language data, and not just in the synthetic exercises common in textbooks (e.g. 'Do the following segments form a natural class?'). It is especially important for practice data sets to include patterns of phonemic contrast and allophonic variation that differ from the patterns in the students' native language and in the language of instruction, so that students can learn to suppress the perceptual filters of the languages they are familiar with.

3.2. ASSESSING STUDENTS' THINKING. It was somewhat dismaying to me as an instructor to observe the protocols in which students produced a solution that likely would have earned full points on an exam, then followed their solution with comments that revealed a conceptual misunderstanding. Like the chemistry students in Cracolice and colleagues' (2008) study, these students had a mismatch between their algorithmic reasoning and their deeper conceptual reasoning. It may well be that their conceptual reasoning developed further throughout the course. But in support of such development, I have modified assignments and test questions to probe the conceptual level in addition to the algorithmic level. Some such conceptual questions are:

- Here are two possible word forms in language X. Would you predict that they will have the same meaning or two different meanings?
- If you are learning language X as an adult, is it more important to learn to produce distinction A [a phonemic contrast] or distinction B [an allophonic difference]? Why?
- If a native speaker of language X is learning English, are they likely to be able to hear the difference between word A and word B? Why?

Such questions are important throughout the course, during class activities and in assignments and practice exercises, in order to support students' conceptual development. It is also essential to ask such questions on exams, to gain an accurate assessment of students' conceptual reasoning in addition to their abilities to follow the steps in a problem-solving algorithm.

4. Future research. This think-aloud study has provided insights into novice students' learning of phonology concepts, and these insights have informed my teaching practice. Since students often also struggle with linguistic reasoning in syntax, I am currently carrying out a parallel study in the syntax portion of the course.

In a fascinating modification of the think-aloud method, Berardi-Coletta and colleagues (1995) compared participants' problem-solving performance in a standard think-aloud condition to a second condition in which the researchers asked metacognitive questions such as 'How are you deciding? How do you know this is a good move?'. The authors showed that participants in the metacognitive condition solved the problems in fewer moves and produced more statements about the problem-solving process itself in contrast to the baseline participants whose statements described superficial features of the problem. They found that participants engage in metacognitive-level thinking about their reasoning 'only when … asked to explain what they are doing and why' (Berardi-Coletta et al. 1995:220) and argued for scaffolding metacognitive processes within problem-solving exercises to support learners' development of metacognition.

A worthwhile follow-up to the present study would investigate what kinds of metacognitive processing questions best support the development of students' scientific reasoning about language. Once we have identified the right kind of questions, we can incorporate them into our learning activities and better guide our students toward thinking like linguists.

REFERENCES

Anderson, Catherine. 2012. A linguistics concept inventory for evaluating teaching and learning. Poster presented at the Canadian Linguistic Association, Waterloo, ON.

Anderson, Charles, and Dai Hounsell. 2007. Knowledge practices: 'Doing the subject' in undergraduate courses. *Curriculum Journal* 18.463–78. DOI: 10.1080/09585170701687910.

Ayres, Paul L. 1993. Why goal-free problems can facilitate learning. *Contemporary Educational Psychology* 18.376–81. DOI: 10.1006/ceps.1993.1027.

Banning, Maggi. 2008. The think aloud approach as an educational tool to develop and assess clinical reasoning in undergraduate students. *Nurse Education Today* 28.8–14. DOI: 10.1016/j.nedt.2007.02.001.

Berardi-Coletta, Bernadette; Linda S. Buyer; Roger L. Dominowski; and Elizabeth R. Rellinger. 1995. Metacognition and problem solving: A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21.205–23. DOI: 10.1037/0278-7393.21.1.205.

Bernstein, Jeffrey L. 2010. Using 'think-alouds' to understand variations in political thinking. *Journal of Political Science Education* 6.49–69. DOI: 10.1080/15512160903467695.

Blackwell, Richard T.; John P. Galassi; Merna Dee Galassi; and Thomas E. Watson. 1985. Are cognitive assessment methods equal? A comparison of think aloud and thought listing. *Cognitive Therapy and Research* 9.399–413. DOI: 10.1007/BF01173089.

Catrambone, Richard. 1998. The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General* 127.355–76. DOI: 10.1037/0096-3445.127.4.355.

Charters, Elizabeth. 2003. The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education Journal* 12.68–82. Online: https://brock.scholarsportal.info/journals/brocked/home/article/view/38.

Chase, William G., and Herbert A. Simon. 1973. Perception in chess. *Cognitive Psychology* 4.55–81. DOI: 10.1016/0010-0285(73)90004-2.

Chi, Michelene T. H.; Miriam Bassok; Matthew W. Lewis; Peter Reimann; and Robert Glaser. 1989. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science* 13.145–82. DOI: 10.1207/s15516709cog1302_1.

Chi, Michelene T. H.; Nicholas de Leeuw; Mei-Hung Chiu; and Christian LaVancher. 1994. Eliciting self-explanations improves understanding. *Cognitive Science* 18.439–77. DOI: 10.1207/s15516709cog1803_3.

Chi, Micheline T. H.; Robert Glaser; and M. J. Farr (eds.) 1988. *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum.

Clement, John J. 1998. Expert novice similarities and instruction using analogies. *International Journal of Science Education* 20.1271–86. DOI: 10.1080/0950069980201007.

Cowan, William, and Jaromira Rakušan. 1987. *Source book for linguistics*. Philadelphia: John Benjamins.

Cracolice, Mark S.; John C. Deming; and Brian Ehlert. 2008. Concept learning versus problem solving: A cognitive difference. *Journal of Chemical Education* 85.873–78. DOI: 10.1021/ed085p873.

Dugger, Katie M.; Eric D. Forsman; Alan B. Franklin; Raymond J. Davis; Gary C. White; Carl J. Schwarz; Kenneth P. Burnham; et al. 2015. The effects of habitat, climate, and Barred Owls on long-term demography of Northern Spotted Owls. *The Condor* 118.57–116. DOI: 10.1650/CONDOR-15-24.1.

Ericsson, K. Anders. 2003. The acquisition of expert performance as problem solving. *Psychology of problem solving*, ed. by Janet E. Davidson and Robert J. Sternberg, 31–84. Cambridge: Cambridge University Press.

Ericsson, K. Anders; Ralf Th. Krampe; and Clemens Tesch-Römer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100. 363–406.

Ericsson, K. Anders, and Herbert A. Simon. 1993. *Protocol analysis: Verbal reports as data*. Revised edn. Cambridge, MA: MIT Press.

Forsberg, Elenita; Kristina Ziegert; Håkan Hult; and Uno Fors. 2014. Clinical reasoning in nursing, a think-aloud study using virtual patients—A base for an innovative assessment. *Nurse Education Today* 34.538–42. DOI: 10.1016/j.nedt.2013.07.010.

Genest, Myles, and Dennis C. Turk. 1981. Think-aloud approaches to cognitive assessment. *Cognitive assessment*, ed. by Thomas V. Merluzzi, Carol R. Glass, and Myles Genest, 233–69. New York: Guilford.

Gentner, Dedre, and Donald R. Gentner. 1983. Flowing waters or teeming crowds: Mental models of electricity. *Mental models*, ed. by Dedre Gentner and Albert L. Stevens, 61–93. Englewood Cliffs, NJ: Lawrence Erlbaum.

Getzels, J. W. 1982. The problem of the problem. *New directions for methodology of social and behavioral science* (Question framing and response consistency 11), ed. by R. Hogarth, 37–49. San Francisco: Jossey-Bass.

Gigerenzer, Gerd, and Ulrich Hoffrage. 1995. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review* 102.684–704.

Halle, Morris, and G. N. Clements. 1983. *Problem book in phonology*. Cambridge, MA: MIT Press.

Kuiper, Koenraad. 2011. Teaching LING101. *Teaching linguistics: Reflections on practice*, ed. by Koenraad Kuiper, 182–88. London: Equinox.

Kuipers, Benjamin, and Jerome P. Kassirer. 1984. Causal reasoning in medicine: Analysis of a protocol. *Cognitive Science* 8.363–85. DOI: 10.1207/s15516709cog0804_3.

Larkin, Jill; John McDermott; Dorothea P. Simon; and Herbert A. Simon. 1980. Expert and novice performance in solving physics problems. *Science* 208.1335–42. DOI: 10.1126/science.208.4450.1335.

Lesgold, Alan. 1988. Problem solving. *The psychology of human thought*, ed. by Robert J. Sternberg and Edward E. Smith, 188–213. Cambridge: Cambridge University Press.

Lesgold, Alan; Harriet Rubinson; Paul Feltovich; Robert Glaser; Dale Klopfer; and Yen Wang. 1988. Expertise in a complex skill: Diagnosing x-ray pictures. In Chi et al. 1988, 311–42.

Lively, Scott E.; John S. Logan; and David B. Pisoni. 1993. Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America* 94. 1242–55.

Marton, Ference. 1986. Phenomenography—A research approach to investigating different understandings of reality. *Journal of Thought* 21.28–49. Online: http://www.jstor.org/stable/42589189.

MAYE, JESSICA; JANET F. WERKER; and LOUANN GERKEN. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82.B101–B111. DOI: 10.1016/S0010-0277(01)00157-3.

MEYER, JAN H. F., and RAY LAND. 2003. Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines. Enhancing Teaching-Learning Environments in Undergraduate Courses Project, Occasional report 4, May 2003. Edinburgh: University of Edinburgh.

MEYER, JAN H. F., and RAY LAND. 2005. Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education* 49.373–88. DOI: 10.1007/s10734-004-6779-5.

MEYER, JAN H. F., and RAY LAND (eds.) 2006. *Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge*. London: Routledge.

MIDDENDORF, JOAN, and DAVID PACE. 2004. Decoding the disciplines: A model for helping students learn disciplinary ways of thinking. *New Directions for Teaching and Learning* 98.1–12. DOI: 10.1002/tl.142.

MILLER-YOUNG, JANICE E. 2013. Calculations and expectations: How engineering students describe three-dimensional forces. *The Canadian Journal for the Scholarship of Teaching and Learning* 4:4. DOI: 10.5206/cjsotl-rcacea.2013.1.4.

NEWELL, ALLEN, and HERBERT A. SIMON. 1972. *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

NURRENBERN, SUSAN C., and MILES PICKERING. 1987. Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education* 64.508–10. DOI: 10.1021/ed064p508.

OWEN, ELIZABETH, and JOHN SWELLER. 1985. What do students learn while solving mathematics problems? *Journal of Educational Psychology* 77.272–84.

PACE, DAVID, and JOAN MIDDENDORF. 2004. *Decoding the disciplines: Helping students learn disciplinary ways of thinking*. (Special issue of *New Directions for Teaching and Learning* 98.) Oxford: Wiley.

PRETZ, JEAN E.; ADAM J. NAPLES; and ROBERT J. STERNBERG. 2003. Recognizing, defining, and representing problems. *Psychology of problem solving*, ed. by Janet E. Davidson and Robert J. Sternberg, 3–30. Cambridge: Cambridge University Press.

REED, STEPHEN K. 1987. A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13.124–39. DOI: 10.1037/0278-7393.13.1.124.

REED, STEPHEN K., and CHERYL A. BOLSTAD. 1991. Use of examples and procedures in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17.753–66. DOI: 10.1037/0278-7393.17.4.753.

TAGLIAMONTE, SALI A., and ALEXANDRA D'ARCY. 2007. Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change* 19.199–217. DOI: 10.1017/S095439450707007X.

TRABASSO, TOM, and JOSEPH P. MAGLIANO. 1996. Conscious understanding during comprehension. *Discourse Processes* 21.255–87. DOI: 10.1080/01638539609544959.

VAN SOMEREN, MAARTEN W.; YVONNE F. BARNARD; and JACOBIJN A. C. SANDBERG. 1994. *The think aloud method: A practical guide to modelling cognitive processes*. London: Academic Press.

VISHTON, PETER. 2005. Using kitchen appliance analogies to improve students' reasoning about neurological results. *Teaching of Psychology* 32.107–9.

WAINWRIGHT, SUSAN FLANNERY, and PATRICIA QUINN MCGINNIS. 2009. Factors that influence the clinical decision-making of rehabilitation professionals in long-term care settings. *Journal of Allied Health* 38.143–51.

WERKER, JANET F., and RICHARD C. TEES. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7.49–63. DOI: 10.1016/S0163-6383(02)00093-0.

WHITNEY, PAUL, and DESIREE BUDD. 1996. Think-aloud protocols and the study of comprehension. *Discourse Processes* 21.341–51. DOI: 10.1080/01638539609544962.

WINEBURG, SAMUEL S. 1991. On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal* 28.495–519. DOI: 10.3102/00028312028003495.

WISMATH, SHELLY; DOUG ORR; and BRUCE MACKAY. 2015. Threshold concepts in the development of problem-solving skills. *Teaching and Learning Inquiry: The ISSOTL Journal* 3.63–73. DOI: 10.2979/teachlearninqu.3.1.63.

Department of Linguistics & Languages
1280 Main St. W.
Hamilton, ON, Canada L8S 4M2
[canders@mcmaster.ca]

[Received 23 December 2015;
revision invited 28 April 2016;
revision received 1 June 2016;
accepted 15 July 2016]