Annotation and Analyses of Temporal Aspects of Spoken Fluency

HEATHER HILTON
Université de Savoie

ABSTRACT

This article presents the methodology adopted for transcribing and quantifying temporal fluency phenomena in a spoken L2 corpus (L2 English, French, and Italian by learners of different proficiency levels). The CHILDES suite is being used for transcription and analysis, and we have adapted the CHAT format in order to code disfluencies as precisely as possible. We briefly present findings for two extreme subgroups in the corpus—our most hesitant and least hesitant learners—and compare the major differences in the temporal structure of the speech of these two learner groups with a native-speaker control group. Implications of these findings for the automatic assessment of spoken fluency will be discussed.

INTRODUCTION

The importance of standardized oral proficiency assessment was recognized in the US back in the 1980s, when considerable effort was invested in calibrating the descriptors of the American Council on the Teaching of Foreign Languages (ACTFL) Proficiency Scale (Byrnes et al., 1986). In Europe, private and government-sponsored organizations have developed batteries of language-specific instruments that include oral proficiency components: the Cambridge ESOL (English for Speakers of Other Languages) Examinations, the Goethe-Institut Zertifikaten, the Instituto Cervantes Diplomas de Español como Lengua Extranjera, and so forth. A centralized European scale of "reference levels"—published relatively recently but resolutely grounded in 1980s communicative theory[1]—places the notion of oral proficiency, and what exactly we mean when we say an L2 speaker is "proficient," squarely at the center of language policy and debate in Europe at the moment. The most recent national syllabus for foreign language teaching in the French secondary schools, for example, is based on this *Common European Framework* (CEF) of reference levels (Council of Europe, 2001), despite the fact that the descriptors for the reference levels have come under criticism as being underspecified (Hulstijn, 2007) and therefore difficult to implement (Alderson, 2007).

The problems that plague all oral proficiency ratings, from the point of view of validity and reliability, are the subjective nature of evaluator assessments and the amount of time needed for face-to-face oral interviews. Tools for automated oral proficiency testing exist (Pearson's *Versant* diagnostic tests) or are under development (Zechner & Bejar, 2006), but there are limitations to what each type of tool can do. The *Versant* tests limit the speaker's output to repetitions or very short-answer output; the long version of the text has a final, open question, but it is unclear (to outside researchers) which parameters of spoken language are automatically scored in the answer. These tests are certainly valid for diagnostic testing

where the phonological accuracy of L2 speech is paramount, but Hincks (2001) found that a precursor to the current tests was unable to measure gains in spoken L2 proficiency after 200 hours of instruction. Zechner, Higgins, and Xi (2007) report a process by which they trained speech recognition software to identify 33% of the words in spoken English-L2 monologues (up from an initial 15%) and conclude that automatic analysis of L2 speech must be based on multicomponent systems. With automatic word recognition in connected L2 speech running so low, we are still relatively far from automatic lexically based proficiency measures.

The object of this article is to present the methodology used to transcribe (manually) and analyze (automatically and semiautomatically) temporal features in a corpus of spoken L1 and L2 productions with an eye to identifying factors that could be incorporated into just such a multicomponent tool for assessing L2 speech. Temporal disfluencies—present in native speech but in relatively stable or predictable quantities—may indeed be a distinguishing characteristic of L2 production, given the lesser degree of automaticity in the processing of L2 language forms (Kormos, 2006). Before we present our transcription system and the analyses that it enables us to generate, we will briefly summarize the model of spoken production and fluency that underpin both. The work described here is only partially automated, but we hope that the findings from our manually constituted oral corpus may contribute concretely to ongoing discussions of how to automate spoken proficiency assessment.

SPOKEN PRODUCTION, FLUENCY, AND DISFLUENCY

Any description of L2 speaking proficiency needs to be grounded in a fully fledged theory of language production (Hulstijn, 2007). Current models of spoken production posit two major processing systems: a semantic system which "map[s] the conceptualization one intends to express onto some linear, relational pattern of lexical items" and a phonological system which "prepare[s] a pattern of articulatory gestures whose execution can be recognized by an interlocutor as the expression of … the underlying conceptualization" (Levelt, 1999, p. 86). Each of these two "core systems" performs several encoding processes: conceptual structuring and the activation of appropriate lemmas (syntactically specified representations in the mental lexicon) in the semantic system and formal (morphological and phonological) encoding and the preparation and execution of articulatory processes in the phonological system. "Self-perception" (p. 88) or monitoring in which we apply reception processes to our own output enables us not only to rectify those encoding errors that occur about once in every 1,000 words produced (Levelt, 1992), but also to adapt our speech to the changing conversational situation.

In our first language, many of the incremental and overlapping processes of lexical retrieval, morphosyntactic encoding, phonological planning, and execution are carried out automatically without the need for attentional effort in the executive component of working memory. The automatic nature of the formal aspects of L1 encoding is certainly what enables all normally constituted individuals to exhibit complete "fluency" when speaking their native language. We will limit ourselves in this study to a narrow (Lennon, 2000, p. 25), temporally defined notion of fluency: fluent production is characterized by a speaking rate of from 130 to 200 words per minute (2-3 words per second); about one third of production time is spent pausing (partly for encoding purposes and partly to enable our interlocutor to process what we are saying), and, according to early "pausology" studies in the 1960s and 1970s, more than two thirds of the pauses in fluent L1 speech are situated at junctures between conceptual or syntactic units— in other words, at utterance or clause boundaries (Goldman-Eisler, 1961, 1968; Hawkins, 1971; O'Connell & Kowal, 1980; Beattie, 1980; Good & Butterworth, 1980; Levelt, 1989). Clinically disfluent L1 speech is characterized by a speech rate of fewer than

50 words per minute (Marshall, 2000); more frequent, longer pauses chop the speech stream up into shorter "runs" that are less coherent from a syntactic or conceptual point of view (Pawley & Syder, 1983). Speech production is considered to have stopped when a hesitation exceeds 3 seconds (Griffiths, 1991); in interactive speech, a conversation partner will tend to intervene once a pause stretches beyond 2.5 seconds (Rieger, 2003). Retracings—repetitions, reformulations, and restarts—often accompany silent or filled pauses and are another sign of encoding difficulties during the speech production process (Kormos, 2006)

In L2 production, of course, the network of automatically available lexical and mor-phophonological representations is limited; we may follow similar procedures to structure concepts and discourse as in our L1,[2] but encoding difficulties can provoke disfluency at every step of the formulation process: a concept may not activate the appropriate L2 lemma; the lemma may not activate appropriate syntactic, morphological, or phonological routines; and/ or the articulatory apparatus may stumble over less well rehearsed segmental or supraseg-mental combinations. The close examination of hesitation structures in L2 speech therefore constitutes a useful tool for identifying which processing components prove most problematic for learners at different levels: "Hesitations are especially useful in showing us where it is easy to move on [in speech production] and where it is difficult" (Chafe, 1980). And, indeed, after a flurry of "pausological" studies of L2 speech in the 1970s and early '80s (Dechert,1980; Deschamps, 1980; Raupach, 1980; Möhle, 1984), followed by a scientific lull, temporal flu-ency indicators have again become prevalent in the investigation of developing L2 speaking skill (Riazantseva, 2001; Towell, 2002; Freed, Segalowitz, & Dewey, 2004; Kormos & Dénez, 2004; Trofimovich & Baker, 2006; Larsen-Freeman, 2006; O'Brian, Segalowitz, Freed, & Col-lentine, 2007; see also Riggenbach, 2000).

THE PAROLE CORPUS: TRANSCRIPTION OF LOWER LEVEL FLUENCY FACTORS

The PAROLE (*PARallèle, Oral en Langue Etrangère* 'parallel oral foreign language') corpus (Hil-ton et al., 2008)—was designed to include samples from learners of three different languages (English, French, and Italian) at different proficiency levels, performing comparable speaking tasks.[3] In addition to his or her contribution to the spoken corpus, each subject completed a battery of tests and questionnaires designed to furnish supplementary data: knowledge of the L2 grammar and lexicon, L2 listening skill, phonological memory capacity, aptitude for grammatical analysis, motivation for L2 learning, and language profile. All participants were young adults (average age 21) enrolled at the Université de Savoie at the time of the data collection process and were paid minimum wage for the 3 hours devoted to the PAROLE proj-ect. Subjects were not pretested for proficiency level but were simply recruited on a volunteer basis from different degree programs for which differing degrees of language proficiency were hypothesized (e.g., language majors entering the university, language majors nearing the end of their university studies, and nonlanguage majors). A corpus of productions by native speakers (NSs) performing the same tasks as our learners has also been compiled and tran-scribed, providing benchmark figures for fluent L1 speech. We are still in the initial phases of analyzing the corpus, so the findings presented here are those indicating the most important or obvious trends and are basically limited to temporal features (for details of the architecture and contents of the PAROLE corpus, see Hilton, 2008c).

This study presents findings from the English and French corpus, 45 learners and 17 NSs in all, and two of the corpus tasks. In each task, the subject summarized for the proj-ect investigator (in a one-on-one interview-type situation) a short video clip immediately after viewing it on a screen that the project investigator could not see. During these tasks, the investigator followed a protocol of relatively limited interactional behavior, attempting to

encourage relaxed, spontaneous production without interrupting or influencing the subject's utterance construction process (Noyau, 2002, p. 38). The artificial nature of the tasks and the limited interaction during the recording sessions were concessions to a data collection process that we hoped would generate more easily transcribable and comparable sets of utterances.

Originally, the PAROLE project had a linguistic focus: we wanted to examine the pho-nological, lexical, morphological, and syntactic characteristics of different L2 proficiency levels and identify cross-linguistic and L2-specific phenomena. The corpus was not designed to elicit or capture particular behavior from a temporal point of view. However, once into the transcrip-tion process, we became aware of the importance of hesitation phenomena in these samples of L2 speech, and, indeed, our analyses to date have focused primarily on lower level fluency features, such as hesitations and retracings, and their possible relationship to different types of errors (Hilton, 2008a, 2008b) and to aspects of syntactic and conceptual planning (Os-borne & Hilton, 2008). PAROLE is being transcribed using the CHILDES (Child Language Data Exchange System) suite of software and transcription conventions (MacWhinney & Spektor, 1995-2008),[4] which was originally designed for the analysis of morphosyntactic and lexical aspects of child language development. In our transcriptions we have attempted to code lower level fluency phenomena according to the CHAT conventions, adapted as explained in the sec-tions that follow. The sections below include a certain amount of detail, which we are present-ing here so that the decisions made by the PAROLE team will be available to other researchers interested in transcribing temporal features. We do not know of any other publically available learner corpora that include exhaustive transcription of such features, and it seems important to specify the rationale for the coding methodology adopted.

Silent Pauses

The first wave of fluency research in the 1950s and 1960s set the minimal length of a pause in the speech stream at 250ms (Goldman-Eisler, 1968), but more recent studies adopt a 200ms cut-off point (Butterworth, 1980), despite the fact that some breaks in the speech stream for purely articulatory reasons may last a bit longer than a fifth of a second. It does indeed appear that pauses under 250ms may be related to processing issues, as in the following example from a fluent learner of English (L1-German), who is not entirely sure which type of household appliance is featured in the video she is summarizing:

(1)   *030: […] they tried to: […] move a: [/] a [/] **#0_203** a fridge↑ **#0_244** o:r something else .
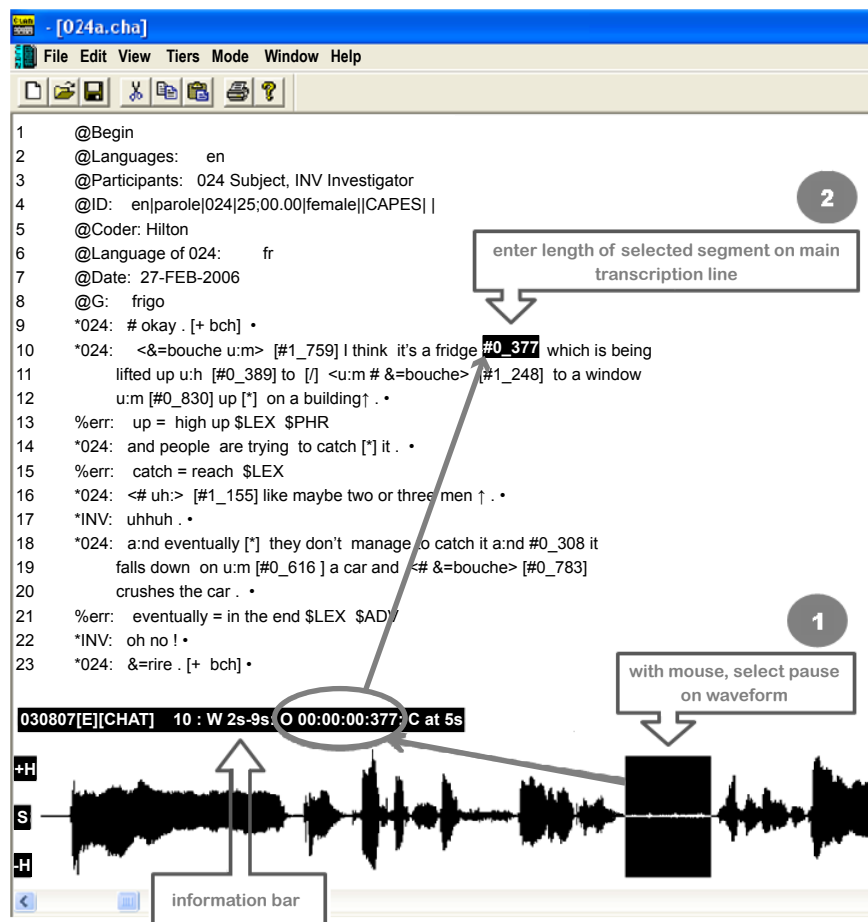
We have therefore coded all hesitations lasting 200ms or more directly on the transcription line in PAROLE, according to the 2007 CHAT convention in which the # symbol indicates a pause, followed by figures indicating the length of the pause (a single underscore replaces the decimal point separating seconds and milliseconds):

(2)   *025: **#0_743** and they were trying to catch something heavy […]
        (fluent learner utterance beginning with a silent pause lasting 743ms)

Pause length (in milliseconds) is easily measured directly on the waveform bar in the "Sonic" coding mode in CLAN (the mode used when linking digital recordings of subject pro-ductions to the CHAT transcriptions). Figure 1 illustrates this timing procedure: the transcriber first selects the section on the waveform bar that corresponds to the silent pause (easily

identified, visually); the black information bar, just above the waveform to the left, gives the precise length of the selected segment (circled in grey in Figure 1). The transcriber can then enter this information directly onto the main transcription line. Throughout the transcription and verification processes, we have taken special care not to overestimate the length of pauses and hesitations in PAROLE, leaving a buffer zone of a few milliseconds at either end of the segment measured, which is not then included in the pause time. In the example shown here, the lingering fricative at the end of "fridge" has been excluded from the pause segment.

Figure 1
Timing Silent Pauses in CLAN

In our coding of hesitations, we have not attempted to distinguish between planning pauses, articulatory pauses, breathing pauses, pauses serving rhetorical functions, and true interruptions to the speech flow, despite the fact that such distinctions are frequently recommended (e.g., Deese, 1980). In fact, any single hesitation may fill more than one function (Rochester, 1973), and it is extremely difficult (if not impossible) to decide on a single function for every pause in the corpus simply by listening to the sound file. In fact, interpreting a speaker's reasons for pausing requires sophisticated analysis and has become one of the major foci in our examination of the PAROLE data.

Filled Pauses

Although filled pauses have certain pragmatic (Clark & Fox Tree, 2002) and discursive functions (Swerts, 1998), they are indeed hesitation phenomena and have all been transcribed in PAROLE: *uh, euh, eh, um,* and *em* (depending on vowel quality and the presence or absence of a bilabial or liquid finish). Filled pauses lasting less than 200ms are simply transcribed where they occur in the speech stream (and not timed). Filled pauses lasting more than 200ms have all been timed (following the same procedure depicted in Figure 1) and the length of the hesitation entered on the main transcription line in square brackets immediately after the transcription of the filler syllable. Drawling of the filler vowel is indicated by the ":" symbol.

(3)   *034: just befo:re **u:m [#0_557]** getting it in to the apartment **u:h [#0_243]** it fell down on a car […] .
       (learner production containing two drawled filled pauses, the first lasting 557ms and the second lasting 243ms)

Filler words such as *well*, *okay*, and *like* in English and *ben* or *enfin* in French have been transcribed but not timed—they are treated as normal words.

Paralinguistic Noises

In addition to silent and filled pauses, breaks in the speech stream can also be filled with a variety of paralinguistic noises, including tongue clacking, sighs of one kind or another, snorts, or sniffs. These sounds have been indiscriminately transcribed in PAROLE as what the CHAT manual calls "simple local events" (MacWhinney 2007, p. 56), using the "&=" symbols and the French word for mouth: "&=bouche" (which must be added to the CLAN depfile). Finger snapping, which occurs occasionally when a subject is looking for a word, is coded "&=snap," and throat clearing is coded "ahem." Loud breathing noises, which appear to characterize the speech of individual subjects, have not been specially coded; they are included with the silent pauses discussed above. Laughter—whether nervous or comic—is coded "&=rire."

Hesitation Groups

Early in the process of transcribing the learner corpus, we were struck by the frequent occurrence of sequences of silent pauses, filled pauses, and paralinguistic noises uninterrupted by the production of linguistic forms (words or attempted words). Despite the fact that fluency research has traditionally considered silent and filled pauses separately (but see Rochester 1973), we have "scoped" these complex hesitation groups on our main transcription line between pointy brackets "< >" (according to CHAT convention) and entered the total duration of the hesitation immediately afterwards in square brackets. So a hesitation group is defined in PAROLE as a sequence of at least two hesitation phenomena (silent pause, filled pause, paralinguistic noise) uninterrupted by an attempt at language production. In the following examples, we see a hesitation group containing two elements that lasts 354ms (example 4a), and a long hesitation group involving several hesitation phenomena, and lasting over 11 seconds (example 4b):

(4a)   *025: and in the end **<# uh> [#0_354]** the: fridge fell #1_103 on a car &=rire .
(4b)   *021: **<u:m # u:h # u:m # &=bouche # u:m uh #> [#11_049]** <I lack the vocabulary> ["] !

Position of Hesitation

In an off-line version of the corpus, the position of each pause (silent or filled) and hesitation group has also been coded, according to three possibilities: at utterance boundaries, clause boundaries, or within a clause (clause internal). It is, of course, possible to code hesitation position in a more detailed fashion, but these three possibilities appear to suffice for the analyses we have conducted. We did, for example, separately code hesitations at noun-phrase or verb-phrase expansion boundaries but have been obliged to assimilate these with clause-boundary hesitations due to the lack of expansions in the production of certain subjects.

Drawls

As already mentioned, "non-phonemic lengthening of syllables" (Raupach, 1980, p. 266) is coded in CHAT with the ":" symbol directly after the letter best representing the lengthened sound (using the same 200ms threshold as for filled and silent pauses). Research into hesitation phenomena in French has concluded that drawling a function word serves the same processing (and pragmatic) needs as a filled pause (Campione & Véronis, n.d.; see also Fox Tree & Clark, 1997). We have therefore timed those parts of drawled function words exceeding 500ms (the median value for a NS hesitation; Hilton, 2008c). Drawled syllables in content words are coded but never timed. Occasionally speakers do drawl content words; there is an example of the word *frigo*: 'fridge' lasting a full second in the French NS corpus (subject N43).

Retracings, Fragments, and Stuttering

In addition to the hesitation phenomena described above, the challenges of spoken production may also give rise to various forms of retracing: repetitions, reformulations, and restarts. In PAROLE, we have slightly adapted the combinations of slash symbols used to code retracings in CHAT (MacWhinney, 2007). A single slash between square brackets "[/]" is used for any simple repetition in which a word (example 5a) or group of words (example 5b) is repeated with no change:

(5a)　　*N47: […] j'ai vu que c' était une sorte **de: [/] de [/] de** frigo […] .
　　　　　(I saw that it was a sort of [/] of [/] of fridge […] .)

(5b)　　*027: […] it's [*] a machine […] to get it up […] **<to the:> [/] to the** room .

As these examples illustrate, drawling is frequently associated with simple repetition; the absence of a drawl in the repeated material is not considered a phonemic reformulation unless the vowel sound also changes in the retracing process (see example 8a).

　　Sublexical fragments and stutters are coded with the "&" symbol (example 6); stuttering is not coded as a repetition unless the fragment is part of a group:

(6)　　*002: […] so we can see a: [#0_620] **&fri frigo↑@n** [*] .

Augmentative duplications (relatively frequent in spoken French) are not coded as retracings since they do not in fact constitute hesitations:

(7)　　*417: […] juste […] **très très** proche [*] […] au [*] bonbon […]
　　　　　([…] just very very close to the candy)

Repetitions involving one change, which we call simple reformulations,[5] are coded "[//]." Linguistic reformulation may involve a change at one of four levels: phonological (example 8a), lexical (example 8b), morphological (example 8c), or syntactic:

(8a)   *N15: […] it was about **a: [//]** u:m [#0_377] **a** crane hoisting a: refrigerator up […] . (initial instance of the drawled determiner pronounced /eɪ/, reformulated as schwa)

(8b)   *034: […] and he wants to: **annoy [//] wind him up** with […] this little chocolate Rolo […] .

(8c)   *020: #0_563 a:nd #0_383 the elephant actually […] **<slap &h> [//] slaps hi:m** #0_493 in the face […]

Semantic reformulations involve the addition or reduction of information; we have not coded them specially, although it would facilitate their retrieval to do so:

(9)   *019: a:nd #0_517 the fridge falls **<on a car> [//] on a green car** […] .

Retracings become restarts when more than one element is changed; the coding symbol "[///]" is used for those restarts in which some element(s)—thematic, syntactic, or lexical—of the initial utterance is maintained:

(10)   *N13: […] (be)cause I guess […] **it won't fit up [///] they […] don't want to take it […] up** the stairs so +/. (thematic continuity, "it" and "up" maintained, new syntactic organization)

The symbol [/-] is used when the restart constitutes a change in utterance structure:

(11)   *N47: […] donc […] j' ai vu un appareil […] qui se:rt **[/-]** […] donc un objet uh en haut […] d' un immeuble .
       (so I saw an apparatus which is used to [/-] well an object at the top of a building)

In this false start, the original syntactic plan for the utterance has been abandoned; the difference with a self-interruption (coded «+//.») being that the speaker has not abandoned the utterance altogether.

Errors

Errors in PAROLE are followed by the symbol "[*]" (as illustrated in example 7 above), and the usual CHAT "scoping" procedure is followed if the error involves more than one word; errors in the first part of a repetition or reformulation are not coded. The nature of the error is indicated with a series of abbreviations (usually three-letter combinations, preceded by the "$" symbol) on a secondary error tier. In PAROLE we have retained five main levels of error (phonological, lexical, morphological, syntactic, and referential/discursive); every error in the corpus is coded for one of these levels at least. For each main error level, a second (and sometimes third) abbreviation may indicate more precisely the type of error (e.g., a morphological error involving verb tense, a syntactic error involving adverb position, a lexical error involving L1 interference, etc.). For a complete list of error types and the abbreviations used, see Hilton (2008c).

Stabilizing the Transcriptions

In order to ensure continuity, the initial transcriptions for the two video summary tasks were completed by a single transcriber for each of the project languages. The transcriptions were then checked by a second transcriber. Intertranscriber agreement was not calculated for various reasons: the coding system used in PAROLE is complex, and mastery of the system develops with time and practice. At the outset of the project, all of the transcribers were using CHILDES for the first time, and our team is too small to generate multiple versions of the same transcription. Rather than attempting to quantify an extremely complicated process, we adopted a collective approach, meeting frequently to discuss specific transcription problems, the use of various symbols, and the most effective way of representing the various language phenomena observed in the corpus. Our interpretation of the CHAT transcription conventions—and, indeed, the conventions themselves—evolved during the transcription process. All coding questions, and any area of disagreement between first and second transcriber, were discussed by at least two other members of the PAROLE team, and the consensus entered in the transcription. The stabilized version of the transcriptions have all been checked four times (by two transcribers at least) with special attention devoted to precision in the timing of hesitations and coherence in the coding of retracings and errors. No log was kept of the time spent on this process, but we conservatively estimate 4 to 5 hours of work on average for the 2 to 3 minutes that each subject took to complete the summary tasks with longer transcription times for the less fluent productions. As in all transcriptions, some errors will persist, but we hope to have kept them to a minimum. Once stabilized, the transcriptions were tagged, using the "mor" and "post" programs in CLAN. This tagging was further disambiguated by hand, which added another 10 hours (total) to the transcription process.

SOME OF OUR FINDINGS TO DATE

The transcription process used in PAROLE, which has only a few automated features, does of course give rise to texts which can be automatically analyzed, using the programs in CLAN or other concordancing software.

General Production Measures

The "mlu" program in CLAN calculates the mean length of utterance (MLU, in words) for each speaker, excluding retraced material, fillers, L1 words, and utterances coded as part of the conversational backchannel. CLAN also calculates two measures of lexical richness: (a) type-token ratio (TTR, which can be limited to a set number of words or "lemmatized" by running the command on the tagged "%mor" tier) and (b) the more sophisticated algorithmic measure of lexical diversity known as D (Malvern & Richards, 1997). Numbers of errors can be tallied, using a key word search for the error code, which also enables the researcher to investigate all of the examples of a certain type of error in the subject productions.

Fluency Indicators

The "timedur" program provided with CLAN outputs a textfile table, listing the duration (in milliseconds) of each utterance that has been matched to the sound file in the Sonic transcription mode. This table can be imported into a spreadsheet, and total speaking time calculated for each participant in the recording. This total can then be used to calculate words per minute (minus fillers and repeated or reformulated material, if the researcher so wishes). A key word

search for the "#" symbol used to code silent pauses, filled pauses, and hesitation groups provides the total number of hesitations produced by each subject and generates a list of the timed hesitation values that can be imported into a spreadsheet and tallied to give the total amount of production time that each speaker spent pausing/hesitating. We can then calculate basic fluency measures: percentage of speaking time spent in hesitation, mean length of hesitation, and a mathematically obtained approximation of mean length of run. Key word searches can be used to generate counts of the different types of retracings produced by each subject, and concordancing software can be used to investigate the sort of language involved in each type of retracing.

Findings

We have not yet completed human ratings of the productions in PAROLE and therefore present findings using the quantitative production and fluency indicators listed above. In order to compare the characteristics of fluent and disfluent learner speech, we have identified two extreme learner subgroups based simply on the percentage of production time spent hesitating. To constitute a disfluent learner subgroup, we took the 13 most hesitant speakers, that is, those who spent over half of their production time hesitating (three NS standard deviations from the most hesitant NS). The fluent learner subgroup was composed of the 13 least hesitant L2 speakers, that is, all of those whose productions fell within the NS hesitation values (less than 35% of production time spent hesitating). The hesitation values found for the two learner groups, the NS group as a whole and the learners as a whole (including the learners who do not fall into the two extreme groups) are presented in Table 1. For the sake of editorial expediency we refer to our extreme subgroups as "fluency" groups, using the term in a strictly temporal sense (and not as a synonym of "proficiency"); there is no doubt more to perceived fluency than mere hesitation time, although Kormos and Dénez (2004) found similar measures of speech rate to be the best predictors of human fluency ratings.

Table 1
"Fluency" Subgroups in PAROLE Based on Percentage of Production Time Spent Hesitating

	Learner corpus (as a whole)	Disfluent learners	Fluent learners	NSs
n =	45	13	13	17
(English corpus)	(33)	(11)	(7)	(9)
(French corpus)	(12)	(2)	(6)	(8)
age	21.4	18.5	23.1	21.3
years of L2 study	8.4	6.5	8.8	–
mean % hesitation time	42.7%	59.9%	28.8%	23.1%
median % hesitation time	40.6%	57.4%	29.6%	21.2%
SD % hesitation time	13.02	7.14	3.87	5.50
minimum % hesitation	20.5%	51.4%	20.5%	13.8%
maximum % hesitation	74.6%	74.6%	34.6%	34.9%

Table 2 provides comparisons of the basic measures obtained from the corpus for these groups and illustrates production differences between fluent and disfluent learners, and L2 learners and NSs performing identical tasks. The last column shows that between-group dif-

ferences (Kruskall-Wallis analyses of variance comparing the disfluent learners, fluent learners, and NSs) are significant for all of the measures except type-token ratio (only marginally different) and is in fact statistically equivalent for the two learner subgroups (U(13, 13) = 72, $p < .50$).[6]

Table 2
Subgroup Comparisons for Corpus Production Measures

Production indicator	All learners ($n = 45$)	Disfluent learners ($n = 13$)	Fluent learners ($n = 13$)	NSs ($n = 17$)	Between-group comparisons
words per minute	105.5	57.4	134.8	183.5	H(2) = 34.60, $p < .0001$
mean length of utterance (MLU, in words)	11.1	8.0	13.0	16.3	H(2) = 30.31, $p < .0001$
mean length of run (MLR, in words)	3.7	2.4	5.1	8.3	H(2) = 32.92, $p < .0001$
type-token ratio (TTR)	.471	.485	.440	.526	H(2) = 7.83, $p < .05$
lexical density (D)	80.91	53.22	103.70	137.73	H(2) = 32.73, $p < .0001$
median length of hesitation	697ms	920ms	572ms	499ms	H(2) = 28.28, $p < .0001$
rate of retracing (per 1000 words)	83.3	128.8	54.9	39.8	H(2) = 16.57, $p < .001$
rate of error (per 1000 words)	93.2	144.8	80.5	10.6	H(2) = 35.21, $p < .0001$

The analyses carried out on the hesitations coded for each of our subjects reveal interesting differences in the temporal structure of learner and native speech. Chi-square goodness-of-fit tests analyzing the distribution of hesitations at the three positions coded in PAROLE (utterance boundaries, clause boundaries, and clause internal) show that this distribution is significantly different for our three subgroups: χ^2(2, 2118) = 97.31, $p < .0001$. NSs hesitate 70% of the time at a discursive or syntactic boundary and only 30% of the time within a clause (this finding is in line with the early "pausology" studies summarized above). Learners' productions exhibit greater proportions of clause-internal hesitations: 54% of the hesitations produced by our fluent learners are situated at a boundary (with proportionally more clause-boundary hesitations than the NSs) and 46% within a clause, whereas over half (56%) of the disfluent learners' hesitations are situated within a clause and only 44% at a boundary. Further chi-square tests reveal that hesitations of differing lengths are distributed differently in learner and native speech (χ^2(15, 3366) = 118.94, $p < .0001$). NS productions contain proportionally more hesitations lasting from 200-600ms (and from 800-900ms), whereas our learners (overall) produced more hesitations lasting from 900ms to over 3 seconds. Comparisons of the subgroups under consideration here show that the disfluent learners produce proportionally more hesitations lasting over 1 second than either the fluent learners or the NSs and that the fluent learners produce proportionally more hesitations in the 500-999ms range than the NSs. Similar chi-square tests analyzing the distribution of the four different types of retracing coded in PAROLE show a marginal difference between the three subgroups: χ^2(4, 735) =11.53, $p < .05$. The disfluent learners produce more simple repetitions and the NSs proportionally more restarts and false starts than either learner group.

DISCUSSION

As the summary of our findings indicates, learner speech is characterized by longer hesitations and numbers of clause internal hesitations which appear to increase as temporal fluency decreases. Clause-internal hesitations have long been attributed to problems with lexical selection (e.g., Goldman-Eisler, 1958), and analyses of PAROLE reported elsewhere (Hilton, 2008b; Osborne & Hilton, 2008) reveal that close to 80% of the disfluent clause-internal hesitations in the PAROLE corpus (those lasting 3 seconds or longer) can indeed be attributed to problems with lexical encoding. Previous psycholinguistic research has also hypothesized that the repetition of a function word is linked to problems with lexical retrieval (Maclay & Osgood, 1959), and our disfluent learners also produced proportionally more simple repetitions than the fluent learners and NSs. The "linear, relational pattern of lexical items" that is automatically generated as we transform ideas into L1 speech (Levelt, 1999, p. 86) is highly problematic for hesitant L2 speakers. Not only is their L2 lexicon much smaller than an L1 lexicon, but the network of associations between lexemes may be relatively impoverished, as illustrated in example 12, where the retrieval of the preposition associated with the verb *fall* requires a concerted effort:

> (12)   *016: […] #0_673 the [/] u:m [#0_400] the fridge u:h [#0_249] fall
> [//] u:h [#0_394] <falls into:> [//] uh falls to [*] the: [/] the car .

The lexical encoding challenges facing L2 speakers is certainly one of the major differences between L1 and L2 production, and we suggest that automatic analyses of learner language focus not only on the quantity and length of hesitations in spoken production, but also on their location in the speech stream.

The Universiteit van Amsterdam is currently completing its large-scale WiSP ("What is Spoken Proficiency?") project designed to investigate the characteristics and correlates of speaking proficiency in L2 Dutch (Hulstijn, Schoonen, de Jong, Steinel, & Florijn, 2004-2008) for the purposes of language assessment and policy (and also to contribute more concrete descriptors to the European reference levels for L2 proficiency assessment). WiSP has given rise to some promising automated measures of the temporal components of spoken proficiency: in particular, programs for the *Praat* software package (Boersma & Weenink, 2009) that automatically total up silences and numbers of syllables in spoken monologues (de Jong & Wempe, 2007). The *Praat* programs, however, cannot distinguish between filled pauses and meaningful syllables, and our manual transcriptions and coding of the hesitations in PAROLE illustrate the importance of filled pauses and complex hesitation groups in learner language. If automatic syllable counts consider filled pauses as syllables, speech rate will be heavily overestimated for certain learners who systematically rely on filled pauses to buy a bit more processing time. PAROLE subjects 003 and 005, for example, produce one filled pause for every two or three words.

An interesting avenue to explore in the automatic analysis of spoken learner language might be a quick manual mark-up of the sound file produced by each subject with a mouse-selection or click-tagging procedure: the examiner could simply tag all hesitations—silent pauses, filled pauses, complex hesitation groups, and even parts of very long drawls. A feature enabling the evaluator to flag clause-internal pauses (any pauses occurring inside utterance or clause boundaries) would then enable the automatic calculation of a variety of fluency measures (coupled with the syllable-counting functions of *Praat*): total speaking time, total hesitation time, amount of speaking time spent hesitating, syllables per minute (excluding filled pauses), mean length of run (in syllables), mean length of hesitation, numbers and clause-internal hesitations, and proportions of hesitations situated at syntactic or discursive boundaries. A speech recognition program's ability to distinguish between syllables and si-

lences within the stretches marked as hesitations would also yield interesting information about the ratio of silent to filled pauses or individual differences in the use of these two different types of hesitation. Such a semiautomatic cross-section of an individual's oral production fluency could be coupled with other (standardized, computerized) measures of his or her lexical and morphosyntactic knowledge of the L2, depending on an institution's evaluation needs (e.g., placement, diagnostic testing, certification, self-assessment, interim feedback, etc.). Semiautomatic fluency detection of this type would also be extremely useful for research purposes: measuring the fallout of different learning situations, practices, or styles and slight changes in the automaticity of L2 use in more tightly controlled experimental situations.

In this study we have presented findings concerning temporal structures in two extreme sets of learner data; we have not yet analyzed hesitation features in the productions by those learners that fall between these two extremes. In order to establish values for hesitation features clearly characterizing L2 proficiency levels, we must first obtain reliable human ratings of the productions in PAROLE; this work is underway for English but has not yet been carried out on the French corpus. In particular, it will be interesting to investigate the distribution of hesitations of more than 1 second and also in the 500-999ms range to establish comparisons with the most and least hesitant learner productions. Another option might be to analyze a set of learner productions that are already rated, such as the spoken samples provided in the European WebCEF project (WebCEF Partners, 2004-2009), using the mark-up procedure outlined above.

We have not addressed, in this article, specific linguistic features in the productions by our learners or language-specific differences within the corpus. From a temporal point of view, we have not found significant differences between the learner productions in English and French or between the two NS groups. Differences certainly exist at the level of individual subjects, and we are currently attempting to identify individual hesitation profiles and examine their relationship to the conceptual and linguistic features of the subjects' productions.

CONCLUSIONS

As the preceding explanations of our coding conventions and corpus analyses illustrate, the rigorous transcription of oral language is a laborious and time-consuming process. Once transcribers have gained a certain expertise, they can proceed efficiently, but the necessary harmonization and checking of the transcriptions requires considerable time and collective effort. The methodology outlined here was developed for the purposes of research, but we hope that a painstaking investigation of spoken language can help identify areas for which automatic measurement can usefully be developed. For future work on the characteristics of spoken production, we will adopt a discourse mark-up program such as EXMARaLDA (Schmidt & Wörner, 2004-2009) or a multimedia annotator such as ELAN (Hellwig, Van Uytvanck, & Hulsbosch, 2002-2009) rather than transcription software designed for textual analysis.

Proponents of task-based oral assessment, such as the authors and defenders of the *European Framework*, will likely be dismayed at the artificial nature of the monologue-type production tasks presented here as a basis for the analysis of spoken language. Whereas the automatic assessment of task accomplishment in listening, reading, and even writing may soon be possible, automatic scoring of interactional spoken language is still a long way off. We draw the reader's attention to data presented by Zechner and Bejar (2006), demonstrating that interrater reliability (human scoring) is lower for integrated speaking tasks (where the speaking task is based on prior reception of spoken or written material) than for monologue-type tasks. If oral proficiency assessment is to do its job, it must be reliable, valid, and fea-

sible. For human raters, for language teachers, for researchers who do not have access to expensive commercially available computerized tests, it is important to understand the most useful ways of breaking oral production into component parts that can be serenely and objectively scored. Initial analyses of our oral L2 corpus indicate that the temporal structure of L2 speech may vary relatively clearly with proficiency level, and we hope that these findings can be incorporated into current attempts at creating a complete battery of descriptors for the rating of spoken proficiency.

NOTES

[1] The average date of publication of the references in the *Common European Framework* bibliography is 1985; the median and mode are both 1987.

[2] However, native speakers of the target language of course can follow different discursive procedures (von Stutterheim, 2003), and we know very little about the acquisition of new conceptual or discursive planning processes (Levelt, 1999).

[3] The three languages chosen for the project reflect the linguistic capacities of the members of our research team; there were no linguistic, social, or psycholinguistic hypotheses underlying the choice of project languages. The findings presented here are taken from the English and French corpus only due to the unfortunate interruption of the transcription and verification of the Italian corpus.

[4] CHILDES includes a set of transcription conventions, called Codes for the Human Analysis of Transcripts (CHAT), and the actual transcription and analysis software, called Computerized Language Analysis (CLAN).

[5] We prefer "reformulation" to "repair," a term that is perhaps used more frequently but which implies that the change produced by the speaker constitutes a correction; this is of course not always the case in L2 production.

[6] Basic TTR is a slightly problematic measure of lexical richness (Malvern & Richards, 1997; Vermeer, 2000), and lemmatized TTR and fixed TTR (type-token ratio calculated on the same number of words for all speakers) calculated for our subjects actually magnify the effect observed here—a higher type-token ratio for the less fluent productions.

REFERENCES

Alderson, J. C. (2007). The CEFR and the need for more research. *Modern Language Journal, 91*, 659-663.

Beattie, G. W. (1980). Encoding units in spontaneous speech: Some implications for the dynamics of conversation. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 131-143). den Haag: Mouton.

Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer* (Version 5.0.46) [Computer program]. Available at http://www.praat.org

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 155-175). London: Academic Press.

Byrnes, H., Child, J., Levinson, N., Lowe, P., Makino, S., Thompson, I., & Walton, A. R. (1986). *ACTFL proficiency guidelines* (1st ed.). Alexandria, VA: American Council on the Teaching of Foreign Languages.

Campione, E., & Véronis, J. (n.d.). *Pauses et hésitations en français spontané*. Retrieved March 21, 2009, from http://74.125.47.132/custom?q=cache:zGB8FDHAw8EJ:sites.univ-provence.fr/veronis/pdf/2004-Campione-JEP.pdf+campione&cd=8&hl=fr&ct=clnk&gl=us&client=google-coop-np

Chafe, W. L. (1980). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 168-180). den Haag: Mouton.

Clark, H. H., & Fox Tree, J. (2002). Using uh and um in spontaneous speaking. *Cognition, 84,* 73-111.

Council of Europe. (2001). *Common European framework of reference for languages*. Cambridge: Cambridge University Press.

de Jong, N. H., & Wempe, T. (2007). Automatic measurement of speech rate in spoken Dutch. *ACLC Working Papers, 2*(2), 51-60. Amsterdam: Universiteit van Amsterdam. Retrieved March 21, 2009, from http://www.hum.uva.nl/aclc/object.cfm/C21601B3-BB4D-4A6A-9C45A56D58D33479/F962E5A9-1321-B0BE-A423747078C3C228

Dechert, H. W. (1980). Pauses and intonation as indicators of verbal planning in second-language speech productions: Two examples from a case study. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 271-285). den Haag: Mouton.

Deese, J. (1980). Pauses, prosody, and the demands of production in language. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 69-84). den Haag: Mouton.

Deschamps, A. (1980). The syntactical distribution of pauses in English spoken as a second language by French students. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 255-262). den Haag: Mouton.

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing 'the' as 'thee' to signal problems in speaking. *Cognition, 62,* 151-167.

Freed, B., Segalowitz, N., & Dewey, D. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition, 26*, 275-301.

Goldman-Eisler, F. (1958). Speech analysis and mental processes. *Language and Speech, 1,* 59-75.

Goldman-Eisler, F. (1961). The distribution of pause duration in speech. *Language and Speech, 4,* 232-237.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* New York: Academic Press.

Good, D. A., & Butterworth, B. L. (1980). Hesitancy as a conversational resource: Some methodological implications. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 145-152). den Haag: Mouton.

Griffiths, R. (1991). Pausological research in an L2 context: A rationale, and review of selected studies. *Applied Linguistics, 12*, 345-362.

Hawkins, P. R. (1971). The syntactic location of hesitation pauses. *Language and Speech, 14,* 277-288.

Hellwig, B., Van Uytvanck, D., & Hulsbosch, M. (2002-2009). EUDICO Linguistic Annotator (ELAN). Nijmegen: Max Planck Institut für Psycholinguistik. Available at http://www.lat-mpi.eu/tools/elan

Hilton, H. E. (2008a). Connaissances, procédures et productions orales en L2. *AILE*, 27, 63-89.

Hilton, H. E. (2008b). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal, 36*, 153-166.

Hilton, H. E. (2008c). Le corpus PAROLE: architecture du corpus et conventions de transcription. *TalkBank*. Pittsburgh: Carnegie Mellon University. Retrieved March 21, 2009, from http://talkbank.org/BilingBank/PAROLE

Hilton, H. E., Osborne, N. J., Derive, M.-J., Suco, N., O'Donnell, J., Rutigliano, S., Billard, S. (2008). *Corpus PAROLE*. Chambéry: Université de Savoie. Available at http://talkbank.org/data/BilingBank/PAROLE/

Hincks, R. (2001). Using speech recognition to evaluate skills in spoken English. *Working Papers in Linguistics, 49.* Lund: Lunds Universitet. Retrieved March 21, 2009, from http://www.speech.kth.se/~hincks/papers/fon01.pdf

Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *Modern Language Journal, 91*, 663-667.

Hulstijn, J. H., Schoonen, R., de Jong, N., Steinel, M., & Florijn, A. (2004-2008). *What is speaking proficiency? (WiSP)*. Amsterdam: Amsterdam Center for Language and Communication, Universiteit van Amsterdam. Available at http://www.hum.uva.nl/wisp

Kormos, J. (2006). *Speech production and second language acquisition.* Mahwah, NJ: Lawrence Erlbaum.

Kormos, J., & Dénez, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32,* 145-164.

Kowal, S., & O'Connell, D. C. (1980). Pausological research at Saint Louis University. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 61-66). den Haag: Mouton.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*, 590-619.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25-42). Ann Arbor, MI: University of Michigan Press.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. In W. J. M. Levelt (Ed.), *Lexical access in speech production* (pp. 1-22). Oxford: Blackwell.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). Oxford: Oxford University Press.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15,* 19-44.

MacWhinney, B. (2007). *The CHILDES Project: Tools for analyzing talk: Vol. 1. Transcription format and programs*. Pittsburgh: Carnegie-Mellon University.

MacWhinney, B., & Spektor, L. (1995-2008). *Child language data exchange system*. Pittsburgh: Carnegie Mellon University. Available at http://childes.psy.cmu.edu

Malvern, D., & Richards, B. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 58-71). Clevedon, UK: Multilingual Matters.

Marshall, R. C. (2000). Speech fluency and aphasia. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 74-88). Ann Arbor, MI: University of Michigan Press.

Möhle, D. (1984). A comparison of the second language speech production of different native speakers. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 26-49). Tübingen: Gunter Narr.

Noyau, C. (2002). Les choix de formulation dans la représentation textuelle d'événements complexes: Gammes de récits. *Journal de la Recherche Scientifique de l'Université de Lomé, 2,* 33-44.

O'Brian, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition, 29*, 557-582.

O'Connell, D., & Kowal, S. (1980). Prospectus for a science of pausology. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 3-10). The Hague: Mouton.

Osborne, J., & Hilton, H. E. (2008, September 12). *Propositional structure and L2 fluency: Findings from a spoken corpus*. Paper presented at the EUROSLA Annual Conference. Aix-en-Provence: Université de Provence.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). London: Longman.

Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 263-270). den Haag: Mouton.

Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*, 497-526.

Rieger, C. L. (2003). Disfluencies and hesitation strategies in oral L2 tests. In R. Ecklund (Ed.), *Proceedings of the 2003 Disfluency in Spontaneous Speech Workshop. Vol. 90: Gothenburg Papers in Theoretical Linguistics* (pp. 41-44). Göteborg, Sweden: Göteborg University.

Riggenbach, H. (Ed.) (2000). *Perspectives on fluency.* Ann Arbor, MI: University of Michigan Press.

Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research, 2,* 51-81.

Schmidt, T., & Wörner, K. (2004-2009). *Extensible markup language for discourse annotation* (*EXMARaLDA*). Hamburg: Universität Hamburg. Available at http://www.exmaralda.org

Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics, 30*, 485-496.

Towell, R. (2002). Relative degrees of fluency: A comparative case study of advanced learners of French. *IRAL, 40*, 117-50.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effects of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*, 1-30.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing, 17*, 65-83.

von Stutterheim, C. (2003). Linguistic structure and information organisation: The case of very advanced learners. In S. Foster-Cohen & S. Pekarek Doehler (Eds.), *EUROSLA Yearbook 2003* (pp. 183-206)*.* Cambridge: Cambridge University Press.

WebCEF Partners (2004-2009). *WebCEF collaborative assessment of oral language proficiency*. Leuven, Belgium: Katholieke Universiteit Leuven. Available at http://www.webcef.eu

Zechner, K., & Bejar, I. (2006). Towards automatic scoring of non-native spontaneous speech. In R. C. Moore (Ed.), *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics* (pp. 216-223). New York: Association of Computational Linguistics.

Zechner, K., Higgins, D., & Xi, X. (2007). SpeechRater™: A construct-driven approach to scoring spontaneous non-native speech. In *1-3 October, Farmington PA. Vol. SLaTE-2007: Proceedings of the 2007 workshop of the international speech communication association (ISCA) special interest group on speech and language technology in education* (pp. 128-131). Pittsburgh, PA: Carnegie-Mellon University.

AUTHOR'S BIODATA

After completing a Ph.D. in French literature and narrative theory at Emory University (1989) and accepting a post as lecturer in English at the Université de Savoie (Chambéry, France), Heather Hilton decided to reorient her research towards what she likes to think of as "applied second language acquisition theory"—thorough grounding of foreign language teaching methodology in the findings of rigorous scientific research into the processes by which individuals acquire and process a foreign language. Her primary research interest is the role of different memory structures in L2 processing. Once her work on PAROLE is finished, she plans to concentrate on the possible links between individual differences in executive short-term memory function, online L2 processing, and long-term L2 learning.

AUTHOR'S ADDRESS

Heather Hilton
Laboratoire LLS, UFR-LLSH
Université de Savoie
73011 Chambéry, France
Phone: (33) 479 758 522
Email: hilton@univ-savoie.fr