Language disintegration in spontaneous speech in Huntington's disease: a more fine-grained analysis

Antonia Tovar[1], A. Garí[2], J. Ruiz-Idiago[3,4], C. Mareca Viladrich[4], E. Pomarol-Clotet[5], Joana Rosselló[6], Wolfram Hinzen[1,5,7]

[1]Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain.
[2] LIMSI, CNRS, Université Paris-Saclay, Orsay, France
[3] Department of Psychiatry and Forensic Medicine. Universitat Autònoma de Barcelona,
Barcelona, Spain
[4] Neuropsychiatry Unit. Hospital Mare de Déu de la Mercè, Barcelona, Spain FIDMAG
Germanes Hospitalàries Research Foundation, Barcelona, Spain
[5] FIDMAG Germanes Hospitalàries Research Foundation, Barcelona, Spain
[6] Department of Catalan Philology and General Linguistics, Universitat de Barcelona,
Barcelona, Spain
[7] ICREA (Catalan Institution for Research and Advanced Studies), Barcelona, Spain

Corresponding author: Wolfram Hinzen

Declarations of interest: none.
Shared first authors: Antonia Tovar, Aina Gari

Highlights

Huntington's disease (HD) involves changes in language, along with primary motor symptoms.
Linguistic change emerged in spontaneous speech prior to the onset of motor symptoms.
They were largely independent of non-linguistic cognitive measures.
Language may be a promising biomarker of disease progression in HD.

Keywords: Huntington's disease, language impairment, grammatical deficits, basal Ganglia

1. Introduction

Huntington's disease (HD) is an autosomal dominant genetic neurodegenerative disease that involves neural death particularly in the striatum (caudate and putamen) (Bano, Zanetti, Mende, & Nicotera, 2011) and causes motor impairments (involuntary movements, chorea). Cognitive and psychiatric symptoms accompany the clinically primary motor symptoms. Cognitive impairments include deficits in executive functions, working and episodic memory, processing speed, and social cognition (Ho et al., 2003; Papoutsi, Labuschagne, Tabrizi, & Stout, 2014; Foroud et al., 1995). In milder forms, these can characterize a prodromal stage of the disease, 15 years or more prior to motor symptoms becoming clinically manifest and crossing diagnostic thresholds (Stout et al., 2011; Paulsen et al. 2014; Bora, Velakoulis, & Walterfang, 2016). As this prodromal period is a critical one for therapies delaying or even preventing symptomatic disease onset, much attention has been devoted to detecting and evaluating the most promising biomarkers that can predict and track disease progression in this phase, including neurocognitive performance in domains such as executive functioning (Paulsen et al., 2014; Wiecki et al., 2016), verbal episodic memory (Solomon et al., 2007), and working memory (Poudel et al., 2015). In the domain of language, too, prodromal changes have been detected, including in the domains of word morphology (regular but not irregular verb and noun inflection: Nemeth et al., 2012), action semantics and sentences with embedded clauses (García et al., 2017; Hinzen et al., 2018). Speech-acoustic aspects such as vowel phonation or speech rate and alterations in steadiness in syllable repetition tasks have also been documented (Vogel, Shirbin, Churchyard, & Stout, 2012; Rusz, Saft, Schlegel, Hoffman, & Skodda, 2014; Skodda, Grönheit, Lukas et al., 2016).

Language impairment is expected from neural atrophy in the basal ganglia, which have been argued to play an important role in non-motor cognitive functions including language (Graybiel, 1995; Ullman et al., 1997; Friederici & Kotz, 2003; Kotz & Schwartze, 2010; Moro et al., 2001). Fronto-striatal circuits may specifically support the building and sequencing of hierarchical structures in language, with phrases embedded in other phrases (Lieberman, 2007), though this process may also depend on more specialized and evolutionary more recent cortical mechanisms (Friederici, 2017). Systematic and detailed behavioural linguistic profiles could

inform debate of the role of the basal ganglia in language, yet are still missing at any phase of the disease. To this purpose, language needs to be assessed as a multi-dimensional construct organized at multiple levels (phonology, morphology, syntax, semantics, discourse), which are re-fitted together into an integrated functional whole. The few behavioural linguistic studies of spontaneous connected speech in HD have typically found a pattern of reduced syntactic complexity, with fewer words and syntactic structures formed in short, simple sentence constructions, more paraphasic and grammatical errors, and sentence truncations (Podoll, Caspary, Lange, & Noth, 1988; Murray & Lenz, 2001; Gordon, & Illes, 1987; Illes, 1989; Chenery, Copland, & Murdoch, 2002; Jensen, Chenery Copland, 2006). These language deficits may form a distinctive signature profile of HD as compared with Parkinson's disease (Murray & Lenz, 2001; Illes, 1989) and people with non-thalamic subcortical lesions (Jensen et al., 2006). Compared to both of these other groups, in particular, HD (at least in early stages) may affect syntactic abilities more than lexical-sematic ones. In later stages of the disease, naming and lexical retrieving difficulties have also been documented in HD (Gordon & Illes, 1987; Illes, 1989; Caine, Bamford, Schiffer, Shoulson, & Levy, 1986), but the origin of these difficulties is unclear and may not relate to semantic memory per se, as opposed to retrieval difficulties and difficulties of visual analysis (Hodges, Salmon, & Butters,1990; 1991). In the case of action words, semantic deficits can characterize prodromal and early symptomatic stages as well (García et al., 2018).

Progress in investigating spontaneous speech production in HD depends on addressing a number of limitations. Thus, sample sizes of the above studies have been small (typically fewer than 12 HD participants), participants with HD at different stages of the disease have often been mixed, and no prodromal cases matched with non-gene carrying neurotypical controls have been included. Moreover, when measuring syntactic complexity, generic measures of complexity (e.g. utterance length or number of embedded clauses, without distinguishing specific forms of embedding, as in Murray & Lenz, 2001; Gordon & Illes, 1987) have typically been used, so that it is unclear which specific aspect of syntactic structuring is compromised. Hinzen et al. (2018) recently set out to address some of these limitations, seeking to profile spontaneous speech

production in HD more comprehensively and to identify those aspects of language structure and function that might differentiate the narrative speech of two groups of identified HD gene-carriers, one prodromal and the other in the early stages of the disease. These groups were compared against each other and that of age- and education-matched neurotypical controls. Speech was elicited through a fairytale retelling task and annotated for a large set of 57 fine-grained linguistic variables (e.g. 'multiple functional word repetition', 'truncation within a word with morpheme integrity preserved', 'hanging determiners', etc.). To create a comprehensive linguistic 'map', these individual variables were then grouped into five broad linguistic 'domains', for which composite measures were computed: 1. *Quantitative* (e.g. number of words produced, mean length of utterance), 2. *Fluency* (e.g. repetitions, pauses, truncations), 3. *Clausal Connectivity* (e.g. use of coordinations such as *and* vs. subordinations such as *(said) that* to connect clauses), 4. *Reference* (use of noun phrases to pick out story characters and maintain topics), and 5. *Concordance* (e.g., marking of grammatical agreement and other morpho-syntactic aspects). Results revealed that narrative speech in early-manifest HD was different in all of these domains relative to the matched controls. Two domains (Reference and Connectivity) showed impairments in pre-manifest HD relative to controls, at a stage of the disease when standardized neuropsychological test profiles were still normal. Scores in the Quantitative but no other domain significantly correlated with the overall Unified HD Rating Scale (UHDRS) motor scores, with working memory scores (Digit Span Backwards; Wechsler, 1981), and with gray matter volume bilaterally in the dorsal basal ganglia (putamen/pallidum). No other domain than Quantitative showed any significant correlations with measures of neurodegeneration. Fluency and Reference correlated with an executive functioning task (the Trail Making Test; Tombaugh, 2004). The remaining two domains, Connectivity and Concordance, did not correlate with any non-linguistic neuropsychological or volumetric measures.

These findings from spontaneous speech production stand in the context of several studies of linguistic comprehension or perception in controlled experimental settings, which have also shown specific linguistic functions to be affected in HD. Sambin et al. (2012) documented this

for aspects of the Binding Theory, i.e. syntactic rules governing co-referentiality between two noun phrases, independently of working memory demands. Teichman et al. (2005) argued for a specific role of the striatum in the application of syntactic and morphological rules, but not lexical knowledge (but see Longworth, Keenan, Barker, Marslen-Wilson, & Tyler, 2008). Teichmann, Dupoux, Kouider, & Bachoud-Lévi (2006) generalized this dissociation to the perception of morphological rules and showed it to be uncorrelated with executive functions. Teichmann, Dupoux, Cesaro, & Bachoud-Lévi (2008) further refined this account through evidence that specific syntactic rules (i.e. syntactic movement), rather than syntactic or combinatorial rules in general, are affected by striatal degeneration; and Teichmann et al. (2008) showed that while syntactic rules are affected over lexical rules, there are impairments in both, traceable to distinct striatal sub-regions and disease stages (see also De Diego-Balaguer et al., 2008).

Together, these findings suggest that language is affected over and above aspects of speech-motor articulation in early HD, and they cast significant doubt on the traditional view that there are no primary language deficits in HD, i.e. these only 'develop secondary to other neurobiological and neuropsychological changes' (Podoll et al., 1988; see also Murray & Lenz, 2001; Gagnon, Barrette, J., & Macoir, 2018). Normal language processing requires and integrates domain-general cognitive functions such as working memory or cognitive control (Just & Carpenter, 1992; Caplan & Waters, 1999; Walker, 1996), hence deficits in these are also expected to bear on language function. However, language deficits documented can concern rather specific linguistic variables; moreover, they are seen in early and even pre-manifest gene-carriers without any other neuropsychological impairment, and there is a lack of correlations with neuropsychological measures in some core aspects of linguistic function (e.g. connectivity). This reinforces the idea of a primary language impairment in HD caused by the neurodegeneration involved. They also strengthen the case for language performance and processing as a potential cognitive biomarker (García et al., 2017; Vogel et al., 2012), in addition to its being an important target for remediation and protective measures, given the importance of language in daily social

functioning (Klasner & Yorkston, 2001; Hamilton et al., 2012; Hartelius, Jonsson, Rickeberg, & Laakso, 2010).

Several questions, however, arise. First, given the methodology of Hinzen et al. (2018), their results reveal little about differences between groups at the level of the fine-grained, non-composite linguistic variables that were factored into the overall domain-level composite scores. This may also have been the reason that these authors failed to find neural correlates of language dysfunction in any except the Quantitative domain. More fine-grained behavioral profiling of linguistic functions is needed to identify language patterns in HD at different stages and to inform future structural and functional neuroimaging studies. These could then also further address the role of the basal ganglia in language processing (Moro et al., 2001; Friederici & Kotz, 2003; Friederici, Steinhauer, & Frisch, 1999), and the link between motor and language functions in the brain more broadly (Lieberman, 2007). A second question concerns the relation and possible interactions between language and cognition in HD. Aphasia-based models of the interface between these two have often stressed the independence of linguistic from general cognitive functioning (Fedorenko & Varley, 2016). However, there is evidence that cognitive decline in putatively nonverbal tasks in fact systematically accompanies language impairment in aphasia (Fonseca, Ferreiras, & Martins., 2016; Baldo Dronkers, Wilkins et al., 2005). In turn, impairment in non-linguistic cognitive domains can contribute to aphasic language performance (Swinney, Zurif, Prather, & Love,1996; Wright, Downey, Gravier, Love, & Shapiro, 2007), stressing the interdependence of language and cognition. Linguistic and cognitive development are closely intertwined as well (Arunachalam & Waxman, 2010), and there is strong evidence for correlations between language performance and performance on standardized ToM tasks in particular (De Villiers, 2007; Paynter & Peterson 2010). Like aphasia and development, HD is an important model for studying this interdependence between language and nonverbal cognition further.

To begin addressing these questions, we had three aims. The first aim was to reproduce the pattern of domain-level results from Hinzen et al. (2018) in an independent cohort with speech samples obtained from different tasks and in more quantity. The second aim was to move from

the composite measures of that study to a more specific and fine-grained level of linguistic analysis with non-composite variables. The third aim was to cast further light on the relation between linguistic performance and cognitive performance as assessed through standardized neuropsychological tests.

2. Materials and Methods

2.1. Participants

This cohort consisted of 20 participants identified as carrying an abnormal polyglutamine expansion in the N-terminal region of the huntingtin protein (Htt) caused by a mutation in the IT15 gene located in chromosome 4, who were matched to 20 neurotypical controls on age, gender, educational background, and (pre-morbid) IQ. Pre-morbid Intelligence Quotient (IQ) was evaluated by the Word Accentuation Test, Spanish version (TAP, Test de Acentuación de Palabras, Gomar, Ortiz-Gil, McKenna et al., 2011). Ten of the 20 HD gene-carriers (which will be referred to below as the 'pre-manifest' HD-group) presented with a score of less than 4 in the Diagnostic Confidence Level (DCL) of the Unified Huntington's Disease Rating Scale (UHDRS). The remaining 10 gene-carriers were at early stages of the disease (henceforth referred to as 'manifest HD'), identified technically by a score between 7-13 on the Total Functional Capacity Scale. All participants were native Spanish speakers. Table 1 summarizes the demographic, genetic and clinical data from the subjects. All participants received the relevant information about the study and the methods and signed an informed consent to participate in this investigation. This informed consent was approved by the ethics committee of the Universitat de Barcelona and the Hospital Mare de Déu de la Mercè (Germanes Hospitalàries).

Table 1: Demographic, genetic and clinical data

	Pre-manifest (N=10)	**Manifest (N=10)**	**Controls (N=20)**
Gender (M/F)	3/7	2/8	5/15
Age (mean/SD)	38.10 (6.82)	48.70 (10.88)	43.30 (9.72)

IQ (mean/SD)	104.6 (5.72)	102.7 (5.87)	105.65 (5.74)
Education in years (mean/SD)	13.5 (3.62)	11.5 (4.06)	12.2 (3.35)
Number of CAG repeats in the larger HTT allele (mean/SD)	41.90 (3.87)	43.40 (3.13)	-
UHDRS TMS* (mean/SD)	1.20 (1.93)	27.10 (13.20)	-
TFC** (mean/SD)	13 (0)	9.50 (1.34)	-

*UHDRS TMS: Unified Huntington Disease Rating Scale Total Motor Score
**TFC: Total Functional Capacity

2.2. Clinical neuropsychological assessment

All participants with HD were evaluated using the motor and functional sections of the Unified Huntington's Disease Rating Scale (UHDRS; The Huntington Study Group, 1996). Their motor performance was described by the total motor scale score (UHDRS-TMS), which was calculated by adding the scores on each of the 31 items of the motor function section of the UHDRS. Each item is rated on a 0 to 4 points scale with 4 indicating the most severe impairment (range 0-124 points). Functional assessment was made using the Total Functional Capacity scale (TFC). Scores on the TFC represent five stages in the neurodegenerative disease process. Lower scores represent greater functional impairment: stage I represents scores of 13–11; stage II, scores of 10–7; stage III, scores of 6–3; stage IV, scores of 2–1; and stage V, a score of 0. The Mini–Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) was used to assess mental status. This is a 30-point questionnaire extensively used in clinical evaluation to measure cognitive impairment. In addition, a cognitive battery was administered consisting of: the Stroop Test (Golden & Freshwater, 1978) assessing inhibition; the Digit Symbol Substitution Test (DSST; Wechsler, 1981) evaluating alternating attention; the Trail Making Test (Tombaugh, 2004) Part A (Trail A), assessing processing speed and sustained visual attention; and Part B (Trail B), evaluating cognitive flexibility. In addition, subtests of the Boston Diagnostic Aphasia Examination test (BDAE; Goodglass & Kaplan, 1972, 1983) were administered in order to assess oral and reading comprehension and naming. Participants also completed one verbal fluency task

(naming of animals). Controls were tested with the same neuropsychological battery to have cognitive indexes characterizing this specific sample and to compare their results with the performance of HD subjects.

2.3. Procedure

Spontaneous speech samples were firstly obtained based on an open, 15-minute interview, in which a list of general questions was specified to structure the conversation (e.g. *Which is your favorite book? Where did you go on holidays?*). Secondly, participants were presented with two video clips, one non-verbal or wordless (1:58 min.) and the other verbal (2:22 min), and asked to retell their story contents to the experimenter. The wordless video showed Mr. Bean faking an accident of falling from a window in order to receive medical attention and care from an attractive nurse. The verbal video presented a discussion between a mother and her daughter in a Chinese restaurant where the waitress casts a spell, causing the mother and the daughter to exchange their bodies the next day. Participants were informed that their speech would be recorded. Speech samples were transcribed and then analyzed utterance by utterance utilizing CLAN (MacWhinney, 2000). The linguistic manual of Hinzen et al. (2018) was used for the linguistic analysis, slightly adapted for the present study; it is added here in Supplementary Materials.[1] Recordings were anonymized.

2.4 Linguistic and statistical analysis

Following the method of Hinzen et al. (2018), a set of 56 individual linguistic variables was chosen for purposes of a comprehensive annotation of spontaneous speech at all levels of linguistic organization, excluding only more peripheral phonetic and articulatory aspects of speech. For analysis purposes, these individual variables were grouped into the same five domains as in Hinzen et al. (2018), capturing different dimensions of linguistic organization. This led to five composite variables, named Fluency, Reference, Connectivity, Concordance and Quantitative. The first four comprise variables capturing different types of errors or anomalies, while the last comprises variables relating to purely quantitative aspects of speech. Specifically,

[1]In the present study, the linguistic variable *Rephrasing* was added.

Fluency was made up of prolongations and repetitions, pauses, and truncations, indexed by the syntactic positions in which these appeared (e.g. pauses between clauses or within noun phrases). **Reference** targeted referential problems inside clauses, such as introducing referents in the discourse that a hearer cannot track, e.g. *Pues no sé por qué. Porque el hermano tenía un piso allí, porque él trabajaba allí* (*I do not know why. Because the brother has an apartment there, because he worked there*, where no such brother has previously been mentioned). **Clausal Connectivity** concerned how clauses were grammatically connected with others, e.g. through coordinations with *and* or subordination with *(said) that…* A characteristic example of a problem of Clausal Connectivity is overuse of coordination in narration, as e.g. in *Y que se lo escuchen más, y que le hagan más caso. Y al final después de insistir tanto y llamar bueno y pues acaba accidentado ¿no?* (*And so that they listen to him and notice him. And at the end, after insisting so much and calling, well, and he ends up injured*). **Concordance** targeted agreement (morphosyntax). Within the domain of Concordance, characteristic examples are agreement failures, like for example in the DP *Las hermanos* (*the brothers*), where the determiner is in feminine gender and the noun in masculine. **Quantitative**, finally, comprised purely quantitative features of speech, such as total number of word/utterances or Mean Length of Utterances/Words. A complete list of variables for each domain is provided in Table 2.[2]

Table 2: List of all variables in each linguistic domain

Fluency	Prolongations (Prol)
	Filled pauses (FilP)
	Lexical word repetition (LWR)
	Single functional word repetition (sFWR)
	Multiple functional word repetition (mFWR)
	Partial functional word repetition (pFWR)

[2]Note that the variables *Infelicitous wildcard* (\$DP), *ambivalence* (±DP), *infelicitous indefinite* (#non-defDP) and *pause within word with morpheme integrity violated* (–WP) were excluded because no participant produced instances of these errors. Two variables were not included in the domain analysis because they did not capture error types, though they were included in the individual analysis: *factive* and *non-factive*, and *Coordination total* and *Subordination total*.

	Partial lexical word repetition (pLWR) Partial repetition of a CP (XPR:CPR) Repetition of a determiner phrase (XPR:DPR) Repetition of a prepositional phrase (XPR:PPR) Repetition of a verb phrase (XPR:VPR) Partial repetition of a phrase (pXPR) Pause between determiner and noun phrase (D-NP.P) Pause between verb and its complement or another clause (V-TP.P:V-CP.P) Pause between clause and tense (C-TP.P) Pause between auxiliary verb and main verb (T-VP.P) Pause between clauses (CP-CP.P) Pause between phrases (XP-YP.P) Pause between a preposition and the following phrase (P-XP.P) Pause between discourse marker and a clause (DM.P) Truncation within a word with morpheme integrity preserved (-W/T) Truncation within a word with morpheme integrity violated (–W/T) Truncation of a phrase after the complementizer (CP/T) Truncation of a quantifier phrase (QP/T) Truncation of a determiner phrase (DP/T) Truncation of a prepositional phrase (PP/T) Truncation of a verb phrase (VP/T) Truncation of a phrase after the auxiliary verb or nexus is uttered (TP/T) Pause within words with morpheme integrity preserved (–WP) Rephrasing (Rephrasing)
Reference	Hanging topic (/top) Abnormal topic shift (#top) Vagueness or lack of topic (0top) Ambivalence (+/–ref) Hanging determiners (/D) Vague referent (VagRef) Definiteness repair (DefRep) Missing determiner (MX:MD) Missing preposition (MX:MP) Failures in temporal reference (refT) Incorrect self-correction of determiners (corXP) Number of mental verbs (v) Paraphasia (Paraphasia)
Connectivity	Missing discourse markers (links) (0D-link) Incorrect discourse marker (D-link) Intrusive parenthetical (#X) Coordination wrong (CRD WRONG) Subordination wrong (SUB WRONG) Coordination total (CRD TOTAL) Subordination total (SUB TOTAL) Failures in consecutio temporum (Tcons.temp)
Concordance	Agreement failure in the auxiliary verb (AgrX:IAgrT) Agreement failure in the main verb (AgrX:IAgrV) Government (GovV) Infelicitous verb (#V)
Quantitative	Utterances (Utterances) Mean length of utterance in morphemes (MLUm)

Total words (WORDS)

Statistical analysis proceeded first at a domain level, then at the level of the fine-grained linguistic variables that made up these domains themselves. Before applying any statistical test, values for every participant were normalized by their number of words or utterances. The choice between normalizing by the total number of words or by the total number of utterances was made based on the nature of the variable. Normalization was not applied to variables in the Quantitative domain. In order to create the composite scores for each domain, variables were first scaled by dividing them by their standard deviation, so as to equate the weight of each variable in the composite score. Composite scores were then obtained by adding the corresponding rescaled variables up. Groups were compared for the composite variables with ANOVA or Kruskal-Wallis tests and the corresponding post hoc tests (Tukey's HSD test and Bonferroni-corrected Dunn test, respectively). ANOVA was used when a normal distribution was present in all groups, as determined by Shapiro-Wilk normality tests. Otherwise, Kruskal-Wallis was applied. Corrections for multiple comparisons were applied to post-hoc group comparisons by means of the Tukey's HSD test itself or with a Bonferroni correction of Dunn test p-values. All p-values of post-hoc tests are reported in their corrected form and can be interpreted with a significance threshold (α) of .05.

Further comparisons were carried out for the individual (non-composite) speech variables. Similarly to the analysis by domain, groups were compared for each variable by means of ANOVA or Kruskal-Wallis tests followed by a post-hoc Tukey's HSD tests or a Dunn test, except in the cases of variables with 50% or more null values (see further below). Corrections for multiple comparisons were again applied only to post-hoc group comparisons. P-values of post-hoc tests are reported already corrected, as described above. Additional corrections by domain were not applied, in order to avoid inflating type II error as well as for comparability with the previous study (Hinzen et al., 2018). We acknowledge the possibility of an increased type I error due to domains being made up of related variables. We provide tables showing Bonferroni-

corrected intercorrelations between variables by domain in the Supplementary Materials (S4-S8). Variables with 50% or more null values were dichotomized in terms of absence and presence, and differences between groups were analyzed with Bonferroni-corrected Fisher's Exact tests. Fisher's test was preferred over χ^2 tests because expected values per group were small.

Comparisons between groups were also explored for the neuropsychological variables by means of ANOVA or Kruskal-Wallis tests. Next, correlation analyses were run between several variables, including speech domain variables and neuropsychological variables. Since the interest was focused on relations between variables in the participants with HD, the group of controls was left out of the correlational analysis. For linguistic variables, correlations in separate groups (pre-manifest and manifest) were also explored. Pearson's or Spearman's coefficients were computed when linearity and monotonicity allowed this. A false discovery rate (FDR) correction was applied to account for multiple comparisons. FDR is the expected proportion of false discoveries amongst all rejected null hypotheses. All reported corrected p-values can be interpreted with a significance threshold of .05.

In all analyses carried out, non-parametric equivalents of parametric tests were applied in cases where variables did not meet the normality assumptions as determined by Shapiro-Wilk tests. Effect sizes of significant results are reported as η^2 or Cramer's V as appropriate. The significance threshold was set at .05 for all tests.[3]

2.5 Reliability analysis

As blindness to medical diagnosis of the participants could only be incompletely ensured, a subset of the sample was re-rated by two independent raters not involved in the study, to check reliability and replicability of rating for the linguistic analysis and the two narrative tasks. Both raters were trained, but as both were linguists, they were largely familiar with the linguistic notions used.

[3] We departed from this general principle only in the intercorrelation tables S4-S8, to make the magnitude of the corrections more perspicuous in this case, where corrections differ from domain to domain.

Transcriptions not involved in the reliability assessment were used for training, which was minimal in both cases. The Intraclass Correlation Coefficient (ICC) was used to measure agreement. In the linguistic analysis, on two out of 56 variables, namely 'Factives' and 'pause within words with morpheme integrity preserved', the ICC could not be calculated due to the lack of variability in the ratings or instances of the relevant variable. Low or null variance does not imply low or null agreement; in fact, most of the ratings in the first of these two variables coincided. We therefore decided to omit these two variables in the agreement analysis. The resulting ICC showed agreement to be very high in general (M=0.950, Median=0.984, SD=0.112). The minimum value was 0.276 in the variable 'agreement failure', very far from the second lowest value (0.774, present in three variables). This variable was omitted in subsequent analyses.

3. Results

Results will be presented in three parts: 1. Domain-level analysis, 2. Individual variables analysis, 3. Neuropsychological results and correlations. Only tables with results of post-hoc comparisons are included here; group-level results can be found in the Supplementary Materials. In the tables, we only include variables for which significant p-values were obtained.

3.1 Domain-level analysis

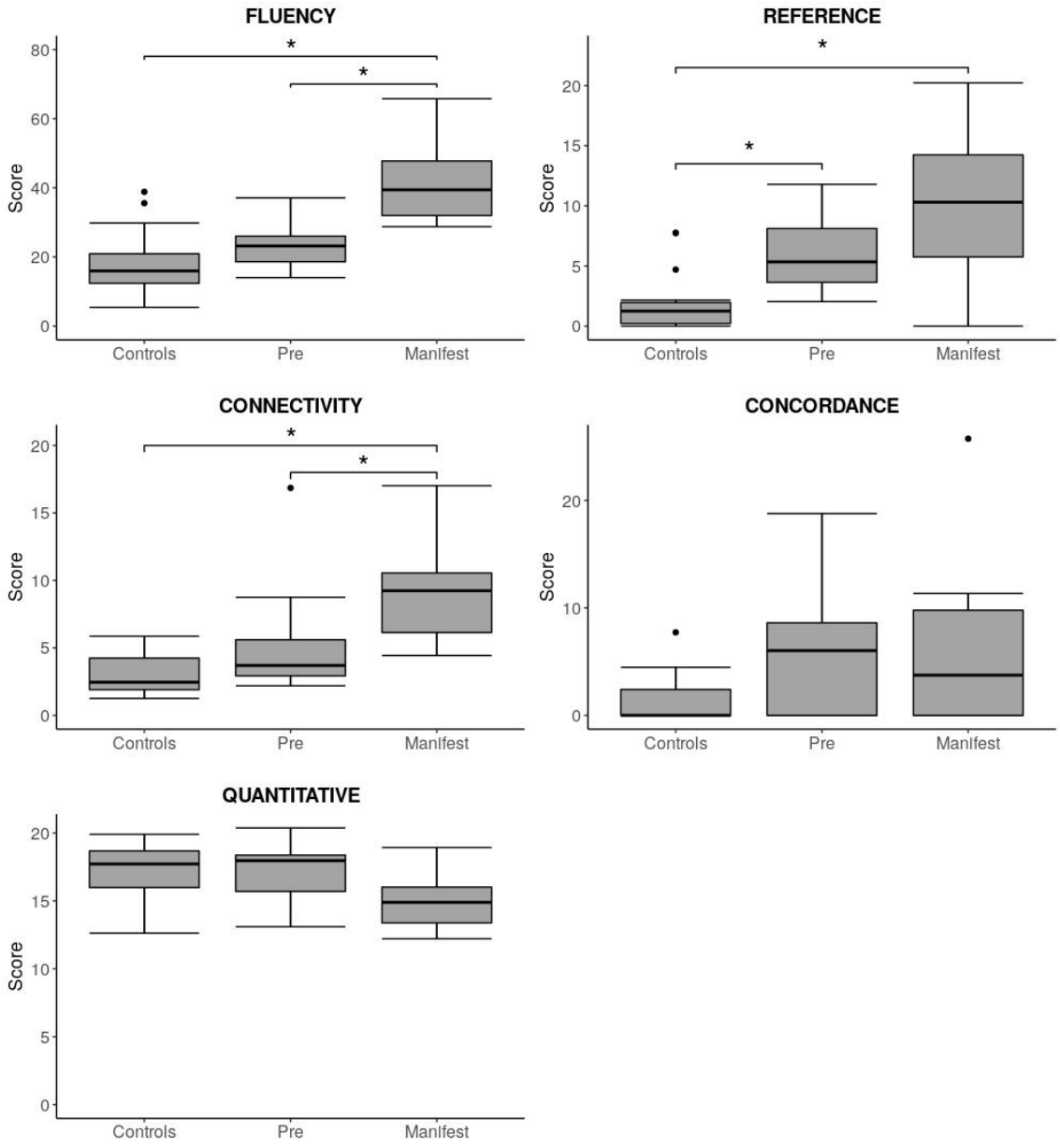
One-way ANOVA and Kruskal–Wallis tests showed there were statistically significant differences between groups in 3 of the error-based domain variables (Fluency: $p<.001$, $\eta^2= 0.530$; Reference: $p<.001$, $\eta^2= 0.488$; Connectivity: $p<.001$, $\eta^2= 0.479$); but not in Concordance: ($p=.068$), nor in Quantitative ($p=.073$) (see also Table S1). Tukey's HSD and Bonferroni-corrected Dunn post hoc tests were respectively applied in order to determine the differences between specific groups showing that there were significant differences in Fluency (control vs. manifest HD: $p<.001$; pre-manifest vs. manifest HD: $p<.001$), Reference (control vs. pre-manifest: $p<.005$; control vs. manifest HD: $p<.001$), and Connectivity (control vs. manifest: $p<.001$; control vs. pre-manifest: $p<.001$; pre-manifest vs. manifest HD: $p<.043$) (see Table 3 and Figure 1 below).

Table 3: Domain-level analysis: post-hoc pairwise group comparisons

Variable (domain)	Mean ± SD	Test	P-values		
			C-P	C-M	P-M
Fluency	C: 17.584 ± 9.195 P: 23.992 ± 7.419 S: 41.108 ± 11.466	Tukey's HSD	.198	< .001*	< .001*
Reference	C: 1.781 ± 2.324 P: 6.011 ± 3.244 S: 10.529 ± 6.483	Dunn's test	.005*	<.001*	.544
Connectivity	C: 3.017 ± 1.424 P: 5.463 ± 4.452 S: 9.089 ± 3.709	Dunn's test	.110	< .001*	.043*

Abbreviations (applying to all result tables): C = controls, P = pre-manifest, M = early manifest. In all result tables, an asterisk (*) indicates significance at $\alpha = 0.05$.

Figure 1: Boxplots for differences between groups in the analysis of linguistic variables grouped into five linguistic domains



3.2. Individual variable analysis

Results of the analysis of non-dichotomized individual variables can be found in Table S2 (general group comparisons) and Table 4 below (pairwise comparisons). [4] Results for dichotomized variables can be found in Table 5 below. In *Fluency*, the following variables were significant in the general group comparisons: Prolongations (Prol, p=.019; η^2=.192), Filled pauses

[4]Note that in three non-dichotomized variables, namely *morphemes*, *pauses between clauses* (CP/T) and *pauses between phrases* (XP-YP.P), Kruskal-Wallis tests did not show a significant difference across the three groups (Table S2), but there were statistically significant differences between controls and participants with manifest HD, as determined by Dunn tests (Table 4).

(FilP, η^2=.266; p=.006), single functional word repetitions (sFWR, p=.005, η^2=.277), Pause between clauses (CP-CP.P,; p=.037, η^2=.169), Pause between discourse markers and/or XP (DM.P, p=.017, η^2=.210), Truncation with morpheme integrity preserved (-W/T, p=.010, η^2=.234), Truncation of DP (DP/T, p=.020, η^2=.201), Truncation of TP (TP/T, p=.003, η^2=.292), Pauses (PAUSES, p=.013, η^2=.221), Rephrasing (Rephrasing, p=.001, η^2=.386), Hanging topic (/top, p=.038, η^2=.168), Coordination wrong (CRD WRONG, , p=.001, η^2=.554), Subordination wrong (SUB WRONG, p=.001, η^2=.527). Pairwise comparisons revealed significant differences between both manifest and pre-manifest HD relative to controls, with both HD groups showing more Filled pauses (FilP) than controls (pre-manifest: p=.005, manifest: p=.039). On the other hand, only pre-manifest HD showed more Prolongations (Prol) relative to controls (p=.028), and only manifest HD showed more Empty pauses (PAUSES) than controls (p=.005). When indexed by syntactic position, these silent lapses occurred in clausal boundary positions, i.e. pauses between or before (as opposed to within) clauses (in formal linguistic terms, either complementizer phrases, CPs, or Tense Phrases, TPs), as reflected in the following variables: Pause between clauses (CP-CP-P: manifest vs. control: p=.016); Pause after a discourse marker, (DM.P (manifest vs. control: p=.013, pre-manifest vs. manifest: p=.022); Pause between V and CP or TP (V-CP/TP.P: manifest vs. control: p=.038, V='verb'), and Pause between full (as opposed to within) phrases (XP-YP.P: manifest vs. control: p=.029). Finally, participants with manifest HD produced more Truncations within words (-W/T, manifest vs. control: p=.007; vs. pre-manifest: p=.017), and within non-clausal phrases (CP/T, manifest vs. control: p=.043, DP/T: manifest vs. control: p=.013, TP/T: manifest vs. control: p=.002). In pre-manifest HD, Single functional word repetitions (sFWR), unlike lexical repetitions, were also increased in relation to controls (p=.002).

In *Reference*, the following variable was significant in the general group comparisons: Hanging topic (/top:, p=.038, η^2=168). This variable also distinguished manifest and pre-manifest HD (p=.028). In the dichotomized variables, both Abnormal topic shift (#top) and Vagueness or lack of topic (-0top) distinguished controls and manifest HD (p=.002, Cramer's V=.595 and

p=.009, V=.557, respectively). The same two groups differed in Ambivalence (+/-ref) (, p=.002, V=.612), Vague referent (VagRef) (p=.012, V=.515), while Definiteness repair (DefRep) (p=.002, V=.596) only distinguished controls from pre-manifest HD.

In *Connectivity*, the general group comparisons revealed significant differences in the variables Total coordination (CRD TOTAL: p=.005, η^2=.248), Total subordination (SUB TOTAL: η^2=.201, p=.016), Coordination wrong (CRD WRONG: p<.001, η^2=.554), Subordination wrong (SUB WRONG: p<.001, η^2=.527), and Intrusive parentheticals (#X: p=.018, η^2=.205). In the pairwise comparisons, total coordination (CRD TOTAL) distinguished controls from both pre-manifest HD (p=.035) and manifest HD (p=.005); while Total subordination (SUB TOTAL) distinguished manifest HD from both controls (p=.043) and pre-manifest HD (p=.018). Manifest HD used both coordinations and subordinations least, i.e. had more isolated sentences with no grammatical connections between them, as mediated through coordinating and subordinating devices. HD groups also misused these patterns of grammatical connectivity between clauses most. Thus, pre-manifest HD misused coordinations (CRD WRONG) relative to controls (p=.005) and manifest HD (p=.017), as well as misusing subordinations (SUB WRONG, pre-manifest HD vs. controls: p=.033). Manifest HD also had more misuses of both coordinations and subordinations in relation to controls: both p<.001.

In the *Quantitative* domain, the following variables were significant in the general group comparisons: number of words (WORDS: p=.049, η^2=.154), mean length of utterance in morphemes (MLUm: p=.018, η^2=.195). In pairwise comparisons, participants with manifest HD produced fewer words than participants with pre-manifest HD (p=.038), and they produced shorter utterances than either participants with pre-manifest HD (p=.037) or controls (p=.025).

In *Concordance*, the variable Agreement failure in the verbal domain (AgrX:IAgrT) was significant between controls and pre-manifest HD (p=.049, V=0.452).

Table 4: Non-dichotomized individual variables: post-hoc pairwise group comparisons

Variable	Mean ± SD	Post hoc test	p-value

			C-P	C-M	P-M
Fluency					
Prolongations (Prol)	C: 0.634 ± 0.255 P: 0.921 ± 0.296 M: 0.856 ± 0.297	Tukey's HSD	.028*	.108	.859
Filled pauses (FilP)	C: 0.078 ± 0.042 P: 0.198 ± 0.132 M: 0.171 ± 0.151	Dunn's test	.005*	.039*	.811
Single functional word repetition (sFWR)	C: 0.007 ± 0.006 P: 0.012 ± 0.003 M: 0.009 ± 0.004	Dunn's test	.002*	.306	.122
Pauses between clauses (CP-CP.P)	C: 0.075 ± 0.073 P: 0.092 ± 0.084 M: 0.173 ± 0.118	Dunn's test	.837	.016*	.133
Pause between XP and YP (XP-YP.P)	C: 0.038 ± 0.047 P: 0.042 ± 0.035 M: 0.118 ± 0.108	Dunn's test	.776	.029*	.214
Pause between discourse markers and/or XP (DM.P)	C: 0.013 ± 0.024 P: 0.017± 0.037 M: 0.091 ± 0.141	Dunn's test	1	.013*	.022*
Truncation with morpheme integrity preserved (-W/T)	C: 0.001 ± 0.001 P: 0.002 ± 0.003 M: 0.004 ± 0.003	Dunn's test	1	.007*	.017*
Truncation of CP (CP/T)	C: 0.014 ± 0.013 P: 0.022 ± 0.021 M: 0.040 ± 0.031	Dunn's test	.767	.043*	.275
Truncation of DP (DP/T)	C: 0.018 ± 0.015 P: 0.020 ± 0.021 M: 0.039 ± 0.016	Dunn's test	1	.013*	0.033*
Truncation of TP (TP/T)	C: 0.022 ± 0.021 P: 0.026 ± 0.019 M: 0.059 ± 0.028	Dunn's test	1	.002*	.018*
Total pauses (PAUSES)	C: 0.159 ± 0.159 P:0.195 ± 0.173 M: 0.462 ± 0.370	Dunn's test	.761	.005*	.077
Rephrasing (Rephrasing)	C: 0.112 ± 0.057 P: 0.121 ± 0.073 M: 0.234 ± 0.081	Tukey's HSD	.944	< .001*	<.001*
Reference					

Hanging topic (/top)	C: 0.012 ± 0.020 P: 0.023 ± 0.017 M: 0.033 ± 0.026	Dunn's test	.125	.028*	.886
Connectivity					
Coordination wrong (CRD WRONG)	C: 0.262 ± 0.101 P: 0.382 ±0.091 M: 0.502 ± 0.075	Tukey's HSD	.005*	< .001*	.017*
Subordination wrong (SUB WRONG)	C: 0.021 ± 0.016 P: 0.083 ± 0.107 M: 0.127 ± 0.078	Dunn's test	.033*	< .001*	.092
Intrusive parenthetical (#X)	C: 0.009 ± 0.016 P: 0.028 ± 0.040 M: 0.037 ± 0.036	Dunn's test	.127	.011*	.605
Total coordination (CRDTOTAL)	C: 0.534 ± 0.137 P: 0.680 ± 0.145 M: 0.460 ± 0.157	Tukey's HSD	.035*	.389	.005*
Total subordination (SUBTOTAL)	C: 0.751 ± 0.253 P: 0.832 ± 0.335 M: 0.495 ± 0.192	Tukey's HSD	.709	.043*	.018*
Quantitative					
Number of words (WORDS)	C: 1787.950 ± 333.750 P: 1841.400 ± 371.345 M: 1438.900 ± 360.179	Dunn's test	.952	.052	.038*
MLU morphemes (MLUm)	C: 19.796 ± 3.398 P: 20.108 ± 3.802 M: 16.149 ± 3.136	Tukey's HSD	.970	.025*	.037*

Table 5: Dichotomized individual variables: pairwise group comparisons
Percentages are of subjects exhibiting instances of the relevant variable in each group. The effect size is an omnibus group effect.

Variables	**Percentages**			**P-values**			**Effect size (Cramer's V)**
	C	P	M	C-P	C-M	P-M	
Fluency							
Pause between V and CP/TP (V-TP.P;V-CP.P)	25%	60%	80%	.297	.038*	1	0.472
Reference							
Abnormal topic shift (#top)	5%	30%	70%	.262	.002*	.328	0.595

Vagueness or lack of topic (-0top)	10%	20%	70%	1	.009*	.192	0.557
Ambivalence (+/- ref)	0%	20%	60%	.284	.002*	.311	0.612
Vague referent (VagRef)	5%	40%	60%	.086	.012*	1	0.515
Definiteness repair (DefRep)	0%	60%	30%	.002*	.081	.678	0.596
Concordance							
Agreement failure in the verbal domain (-AgrX:IAgrT)	5%	50%	30%	.049*	.262	1	0.452

3.3 Neuropsychological variables and correlations

Results revealed that there were statistically significant differences between groups in all but two of the neuropsychological variables (no differences were found between manifest and pre-manifest groups for the BDAE subtest of Reading Comprehension (sentences) and between all the groups for the TAP/NART); see Tables S3 (for results at the group level) and Table 6 (pairwise comparisons). However, the manifest HD group crossed the threshold for clinical impairment (at a 'moderate' level) only in three neuropsychological tasks, namely Stroop (Reading and Denomination), Trail Making Test (TMT, part A and B), and the Digit Symbol Substitution Test.

Table 6: Neuropsychological variables: Post-hoc pairwise group comparisons

Variable	**Mean ± SD**	**Post hoc test**	**p-value**		
			C-P	C-M	P-M
Boston Listening Comprehension test: commands	C: 10.0± 0.00 P: 9.9 ± 0.32 M: 9.4 ± 0.70	Dunn's test	.731	< .001*	.019*
Boston Listening Comprehension test: Complex Ideative Material	C: 6.0 ± 0.00 P: 5.5 ± 0.71 M: 4.3 ± 1.16	Dunn's test	.086	< .001*	.025*
Boston Naming test: Naming Response (questions)	C: 10.0 ± 0.00 P: 10.0 ± 0.00 M: 9.5 ± 0.53	Dunn's test	1	< .001*	.001*

Boston Naming test: Visual Confrontation (images)	C: 14.3 ± 0.66 P: 14.0 ± 1.33 M: 11.3 ± 1.64	Dunn's test	1	<.001*	.001*
Boston Reading + Comprehension	C: 5.0 ± 0.00 P: 5.0± 0.00 M: 4.6 ± 0.70	Dunn's test	1	.005*	.018*
Boston Reading Comprehension: sentences	C: 3.0 ± 0.22 P: 2.8 ± 0.42 M: 2.3 ± 0.95	Dunn's test	.529	.005*	.128
Verbal Fluency test (animal naming)	C: 26.5 ± 6.25 P: 24.1 ± 6.24 M: 17.4 ± 5.38	Tukey's HSD	.566	.001*	.046*
Stroop Reading test	C: 111.3 ± 12.95 P: 99.4 ± 20.97 M: 60.4 ± 15.94	Tukey's HSD	.146	< .001*	<.001*
Stroop Denomination test	C: 77.2 ± 10.09 P: 69.9 ± 12.34 M: 38.4 ± 13.29	Tukey's HSD	.242	< .001*	<.001*
Stroop Interference test	C: 43.5 ± 9.03 P: 43.5 ± 9.88 M: 24.2 ± 9.65	Tukey's HSD	.999	< .001*	< .001*
Trail Making test Part A	C: 36.9 ± 34.78 P: 33.0 ± 9.55 M: 84.0 ± 48.01	Dunn's test	.597	<.001*	.012*
Trail Making test Part B	C: 92.1 ± 61.07 P: 94.9 ± 47.91 M: 160.5 ± 89.43	Dunn's test	1	.022*	.089
Digit Symbol Substitution Test	C: 50.6 ± 14.52 P: 51.3 ± 9.78 M: 26.3 ± 11.66	Dunn's test	1	< .001*	< .001*
Mini–Mental State Examination	C: 29.8 ± 0.52 P: 28.6 ± 1.58 M: 26.5 ± 2.27	Dunn's test	.042*	< .001*	.039*

After FDR correction, only two significant correlations between linguistic and neuropsychological domains remained: (i) between TMT-A and the domain of Fluency (p=<.001, r=0.80), and (ii) between TMT-A and the domain of Reference (p=.041, r=0.611).

4. Discussion

Results of the present study showed that, at the domain-level, groups differed in the domains of Fluency, Reference and Connectivity, but neither Quantitative nor Concordance. A progressive decline from controls to pre-manifest to manifest HD was seen in all of the three former domains, with significance thresholds crossed in the comparison of manifest HD and controls in all three cases, while pre-manifest HD differed from controls only in Reference. This pattern exhibits important overlaps but also differences with the previous study of Hinzen et al. (2018), where significant differences between manifest HD and controls were seen in both Quantitative and Concordance as well. Results in Fluency, Connectivity and Reference, on the other hand, were broadly similar. Specifically, they were identical in terms of comparisons between manifest HD and controls, and like in the previous study, also pre-manifest HD differed from controls in Reference. However, in Connectivity they only differed from controls in the previous but not the present study.

A comparison of the demographics of the two samples involved in these two studies, which were recruited in the same region, revealed no significant differences in age, gender, or education. Difference in the results of the two studies suggest that purely quantitative or else formal grammatical measures (e.g. number or length of utterances or grammatical agreement) may be less sensitive, at a domain-level, in capturing the neuropathology in question at a linguistic level: measures in the domains of fluency, connectivity and reference may reveal language decline more reliably and earlier in the disease process. Note that, on the other hand, some of the individual linguistic variables within Quantitative and Concordance showed significant group differences also in the present study. In particular, manifest HD differed from both pre-manifest and controls in the Mean Length of Utterance. A larger sample size may have shown significant group differences in Quantitative at the domain level as well. As for the loss of significance in Connectivity when comparing pre-manifest HD and controls, this difference may in part be due to a difference in how relevant variables were normalized in both studies (in particular, anomalous uses of coordinations and subordinations were normalized relative to total coordinations and subordinations, respectively, in the present study).

This combined outcome from two studies and independent samples and different tasks used for elicitation, provides renewed support that neurodegeneration in HD affects core domains of language functioning in spontaneous speech as well, from the pre-manifest stage. This is in line with reports of neurodegeneration in HD in language-relevant subcortical areas long before clinical symptoms are seen (Aylward et al., 2012; Bano et al., 2011). In the present study, moreover, this again occurred when neuropsychological tests revealed no decline in pre-manifest gene carriers except in the case of the MMSE (see Table 6). Even in manifest HD, MMSE scores were not in the range of impairment and turned out not to correlate with any linguistic variables or domains. We do not interpret language decline ahead of motor symptomatology as suggesting that language function is unrelated to motor function. Rather, motor deficits may be too subtle at the pre-manifest stage to show in domains other than language, which is the most complex and rapid motor action that humans perform (Lipski et al., 2017; Simonyan & Fuertinger, 2015; Simonyan, Ackermann, Chang, & Greenlee, 2016). This underlines the potential role of language as a clinical marker of disease progression (García et al., 2017; Vogel et al., 2012) and calls for longitudinal studies.

Our second aim was to move the domain-level analysis of Hinzen et al. (2018) down to the level of individual variables. Here a telling pattern arose in the domain of Fluency, where participants with manifest but not pre-manifest HD produced more empty pauses (‘speech left blank’, without fillers) than controls (Table 4). Pauses, in the definition of Silverman (1973), are ‘intermittent feedback delay operations, allowing the momentum of semi-automatic speech generation to be halted while information is processed for the appropriate planning of subsequent utterances’. This suggests that differences in such planning are not yet visible at the pre-manifest stage. Pauses, however, can also be ‘filled’, where filling a pause suggests awareness of the break in the flow of speech, along with interpersonal social signaling that the flow of thought continues. While both HD groups had more filled pauses than controls, only pre-manifest HD showed more prolongations and repetitions in relation to controls. These, too, can be ways of bridging a gap and manifesting awareness of its existence. Further in line with this pattern, *definiteness repair*

(DefRep: self-correction of anomalously introduced referents), also indicative of insight into communication failure, was only seen in the pre-manifest group, but not the manifest one. In short, while manifest patients present 'gaps' in their speech (pauses and truncations of utterances and words), pre-manifest HD tend to fill these gaps using prolongations and repetitions.

Importantly, such breakage patterns showed up along clausal boundaries, suggesting the importance of indexing dysfluency patterns by the syntactic positions in which they occur. Clauses are units of structure where relatively complete units of thoughts are encoded. It would be pauses within simple phrases, such as between an article and a noun, or truncations of them, which would point to a problem at the level of lexical retrieval. This pattern, which has been documented for spontaneous speech in both the cases of Alzheimer's disease (Gayraud, Lee, & Barkat-Defradas, 2011) and post-stroke aphasia (Angelopoulou et al., 2018), was not observed here. Instead, the pattern points to a problem in configuring thought-sized units, i.e. units of structure in language encapsulating a complete thought. Further supportive of this conclusion against a specifically lexical problem, word repetition patterns were confined to repetitions of grammatical function words, not lexical items. These results can be contextualized against those of Vogel et al. (2012) on fluency patterns in people with manifest HD, who differed from both controls and a pre-manifest group in speech rate (syllables/total signal time), total speech time, and total silence time, with manifest HD having a lower speech rate and higher total silence time. These authors, however, did not index pauses by syntactic position, nor distinguished empty and filled pauses. As for the pattern of prolongations, fillers, repetitions, and repairs seen in pre-manifest HD, we tentatively interpret this as reflecting 'adaptive strategies' in the sense of Illes (1989: p. 636), i.e. coping strategies in the face of a functional deficit in language. This functional deficit is centered on the construction of appropriate units of language for purposes of thought and reference.

Results in Connectivity cohere with the significance of the clausal boundary in HD just noted. Grammar across human languages avails us of two key ways in which clauses can be combined: one clause can be embedded in another, in which case one is *sub*ordinated to the other;

or they can be *co*ordinated, in which case they are both grammatically independent and the relation between them is symmetric (i.e., none is subordinated under the other). With this difference goes a difference in the thought expressed: a sentence with a subordinated clause embedded under a verb will ipso facto represent how someone represents the world (what he thinks, says, believes, or wants), and hence express a meta-representation. Failure to use subordinations will make it more difficult to express meta-representations, i.e. thoughts about thoughts, and hence reasoning about mental states (ToM). In line with this, clausal embedding of the subordinating type has been widely argued to be a potential mechanism for accomplishing classical ToM tasks (Paynter & Peterson, 2010; Astington & Jenkins, 1999; Steele, Joseph, & Tager-Flusberg 2003). Key group differences that emerged in this study regarding subordination and coordination patterns could thus be cognitively significant, manifesting difficulties in reasoning about mental states. In particular, participants with manifest HD used both of coordination and subordination less than pre-manifest HD, which suggests that, as the diagnostically criterial motor symptoms emerge, grammar also shifts in its organization, becoming more mono-clausal or grammatically unconnected. Moreover, manifest HD had more anomalous *uses* of coordinations (CRD WRONG) in relation to both pre-manifest HD and controls, and the pre-manifest group produced more wrong coordinations and subordinations than controls. This finding calls for studies in which independently noted impairments in ToM in HD as assessed by standardized ToM tests (Brüne, Blank, Witthaus, & Saft, 2011; Eddy, Mahalingappa, & Rickards 2012; Saft et al., 2013; Adenzato, & Poletti, 2013; Bora et al., 2016), would be correlated with language measures directly, and specifically with coordinating and subordinating clause types. It also underlines the need to differentiate earlier composite measures of syntactic complexity (e.g. Illes, 1989; Murray & Lenz, 2001), in which different forms of syntactic complexity are often amalgamated into a single overall measure of complexity. Utterances with coordinated or subordinated clauses are both ‘complex’, yet very different kinds of complexity, corresponding to different cognitive mechanisms and types of thoughts expressed, are involved.

Results in the Reference domain suggest that language decline not only shows in how clauses are combined and the loss of semantic richness and complexity resulting from this, but also in the use of language for purposes of reference, which relates to discourse. The main shifts here, at the level of individual variables, were seen in manifest, but not pre-manifest HD, through a pattern of abnormal topic shifting, setting up a topic without pursuing it (truncated topics), or vagueness and ambivalence of reference, which is still noted and thus 'repaired' only in the pre-manifest group ('definiteness repair' mentioned above).

Turning to our third aim, all of the neuropsychological measures showed significant differences between controls and the manifest HD group, while only the MMSE showed a difference between controls and pre-manifest HD (Table 6). Only the manifest HD group ever crossed clinical thresholds to cognitive impairment, though in no case, impairment was severe. Moreover, they only did so in three measures, namely the Stroop (only Reading and Denomination), Trail Making Test, and Digit Symbol Substitution Test (Table 6). This suggests problems of attention, working memory, and executive functioning, as well as a potential difficulty with visual processing (since all of these tasks are administered via visual stimuli), though performance on other, also visually based tests was close to normal (e.g. Stroop Interference or BDAE Boston Visual Confrontation). However, only one of these tests (the Trail Making Test Part A) correlated with the linguistic domains (Fluency and Reference). The correlation with Fluency would make sense in light of reductions in processing speed which would affect both speech Fluency and the Trail Making Test. Whether the same link explains the correlations with Reference is less clear. Whether the same link could explain the correlations with Reference is less clear, though post hoc analysis to clarify this point revealed that the domains of Fluency and Reference correlated with each other ($p=.041$, $r=0.606$). As an anonymous referee notes, the Trail Making Test Part B is more challenging cognitively and requires processing speed along with attention and working memory. But it did not correlate with either Fluency or Reference, making the previous correlations difficult to interpret. Unfortunately,

direct measures of working memory were not available in the present study, but they correlated only with the Quantitative domain in the previous study of Hinzen et al. (2018).

It is also noteworthy that, although the manifest group differed from both other groups in language tests designed for patients with aphasia (BDAE), performance on these tests was still generally high; and it did not correlate with the linguistic measures introduced in our study. Language is a complex domain that can disintegrate in a large number of different ways and at different levels: language impairment across clinical groups will rarely be the same. This has the important clinical implication that language tests that can detect and assess language patterns in HD should be devised. In this regard, the present results suggest that patterns of linguistic impairment in HD cut across the traditional linguistic levels of 'syntax' and 'semantics', so that this traditional divide would have been unlikely to capture the clinical patterns seen. Indeed, these level descriptors have become problematic within linguistic theory itself (Wiltschko, 2018), and may be of questionable utility clinically.

Apart from correlations with working memory and direct ToM tests, future work calls for replications of our findings in languages other than Spanish and for linking them to patterns of neural degeneration. Hinzen et al. (2018) failed in identifying structural neural correlates for domain-level linguistic impairments except in the Quantitative domain. Although data about neural atrophy were not available in the present study, the significance of clausal connectivity both at the level of Fluency and Connectivity documented here informs the debate on the significance of the striatum and frontal-striatal loops for syntactic structuring, recursion, and the 'chunking' of linguistic information into clausal informational units (Bornkessel & Schlesewsky, 2013; Graybiel, 1995; Lieberman, 2007) under temporal constraints (Kotz & Schwarze, 2010). Clausal embedding would be a particularly useful focus in future functional neuroimaging studies.

In sum, this study has provided further support for language degeneration in early and pre-manifest HD and contributes to a more fine-grained and differentiated profile of the linguistic phenotype of this disease. In HD, language changes precede other cognitive and motor impairment. These clearly lie outside of the speech-articulatory domain, in core domains of

grammatical organization, and they are not easily accounted for by non-linguistic cognitive impairment, whether occurring in participants with pre-manifest or manifest HD. As language capacities fundamentally impact on communicative abilities, careful attention should be devoted to their early detection, to clinical linguistic tests appropriate for this population, and to protective treatments. With regard to the potential role of language as a marker of disease progression, language has already shown distinctive signature profiles and potential as a predictive and diagnostic measure in a number of other neuropathologies, including Alzheimer's disease (Ahmed, Haigh, de Jager, & Garrard, 2013), Parkinson's disease (García et al., 2017), autism (Eyler Pierce, & Courchesne, 2012; Lombardo et al., 2015), and schizophrenia (Bedi et al., 2015; Rosenstein Foltz, DeLisi, & Elvevåg,, 2015; Cokal et al., 2018). As language disintegration systematically differs across all of these neuropathologies, further comparative work should systematically investigate the sensitivity and specificity of language as a marker of disease progression.

Author statement

JR, WH, and AT conceptualised the study; JRI and CMV supervised the clinical aspects of this study; AT ran the experiment and provided the linguistic data analysis; AG advised on the data analysis plan and carried out the statistical analysis; JRI, CMV and EPC advised on methodology and organised patient recruitment; WH, AG, and AT wrote the original draft; JRI and AG commented on it. EPC and WH provided funding.

Declarations of interest: none.

Funding: This research was supported by the Ministerio de Economía y Competitividad (MINECO, Spanish Government), grant FFI2013-40526P, and the Generalitat de Catalunya, Grant SGR-1265, both to W.H.

Acknowledgements: We are grateful to the patients and their families for their participation and thank Catalina Morey and Alba Cerrudo for help with the independent rating of a subsample.

References

Adenzato, M., & Poletti, M. (2013). Theory of mind abilities in neurodegenerative diseases: an update and a call to introduce mentalizing tasks in standard neuropsychological assessments. *Clinical Neuropsychiatry*, 10(5), 226–234.

Ahmed, S., Haigh, A.M.F., de Jager, C.A., & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, 136:12, 3727–37.

Angelopoulou, G., Kasselimis, D., Makrydakis, G., Varkanitsa, M., Roussos, P., Goutsos, D., Evdokimidis, I., & Potagas, C. (2018). Silent pauses in aphasia. *Neuropsychologia*114:41–9.

Arunachalam, S., & Waxman, S. R. (2010). Language and conceptual development. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 548–558.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental psychology*, 35(5), 1311.

Aylward, E.H., Liu D., Nopoulos, P. C., Ross, C. A., Pierson, R. K., Mills, J. A, Long, J.D., & Paulsen, J.S. (2012). Striatal volume contributes to the prediction of onset of Huntington disease in incident cases. *Biol Psychiatry*, 71, 822-828.

Azambuja. M.J, Radanovic, M, Haddad, M.S, et al. (2012). Language impairment in Huntington's disease. *ArqNeuropsiquiatr*. 70:410–415.

Baldo, J. V., Dronkers, N. F., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language*, *92*(3), 240–250.

Bano, D., Zanetti, F., Mende, Y., & Nicotera, P. (2011). Neurodegenerative processes in Huntington's disease. *Cell Death & Disease*, *2*(11), e228–. http://doi.org/10.1038/cddis.2011.112.

Bedi, G., Carrillo, F., Cecchi, G.A., Fernández Slezak, D., Sigman, M., Mota, N., Ribeiro, S., Javitt, D., Copelli, M., & Corcoran, C. H. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *Schizophrenia*, *1*, 15030.

Benton, A.L, & Hamsher, K. *Multilingual Aphasia Examination manual*. University of Iowa; Iowa City: 1976.

Bora, E., Velakoulis, D., &Walterfang, M. (2016). Social cognition in Huntington's disease: A meta-analysis. *Behavioural brain research*, 297, 131-140.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: a new dorsal–ventral stream model of sentence comprehension. *Brain and language*, 125(1), 60-76.

Brüne, M., Blank, K., Witthaus, H., & Saft, C. (2011). "Theory of mind" is impaired in Huntington's disease. *Movement Disorders: Official Journal of the Movement Disorder Society*, 26(4), 671–8. http://doi.org/10.1002/mds.23494

Caine, E. D., Bamford, K. A., Schiffer, R. B., Shoulson, I., & Levy, S. (1986). A controlled neuropsychological comparison of Huntington's disease and multiple sclerosis. *Archives of Neurology*, *43*(3), 249-254.

Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behav. Brain Sci.* 22, 77–94.

Chenery, H. J., Copland, D. A., & Murdoch, B. E. (2002). Complex language functions and subcortical mechanisms: evidence from Huntington's disease and patients with non-thalamic subcortical lesions. *International Journal of Language and Communication Disorders*, 37(4), 459–474.

Cokal, D., Sevilla, G., Jones, W. S., Zimmerer, V., Deamer, F., Douglas, M., Spencer, H., Turkington, D., Ferrier, N., Varley, R., Watson, S., & Hinzen, W. 2018. The language profile of formal thought disorder. *Schizophrenia* (www.nature.com/npjschz/), https://rdcu.be/7aQi.

De Diego-Balaguer, R., Couette, M., Dolbeau, G., Dürr, A., Youssov, K., &Bachoud-Lévi, A. C. (2008). Striatal degeneration impairs language learning: evidence from Huntington's disease. *Brain*, 131(11), 2870-2881.

DeVilliers, J. (2014). Which kind of concepts need language? *Language sciences* 46, 100-114.

Eddy, C. M., Mahalingappa, S., & Rickards, H. E. (2012). Is Huntington's disease associated with deficits in theory of mind?. *Acta Neurologica Scandinavica*, 126(6), 376-383.

Eyler, L.T., Pierce, K., & Courchesne, E. (2012). A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain: a journal of neurology, 135 Pt 3*, 949-60.

Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369*(1), 132-153.

Fisher, N., Happé, F., & Dunn, J. (2005). The relationship between vocabulary, grammar, and false belief task performance in children with autistic spectrum disorders and children with moderate learning difficulties. *Journal of child Psychology and Psychiatry*, 46(4), 409-419.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3), 189-198.

Fonseca, J., Ferreira, J. J., & Martins, I. P (2017). Cognitive performance in aphasia due to stroke: A systematic review. *Int J Disabil Hum Dev*16:127–139.

Foroud, T., Siemers, E., Kleindorder, D., Bill, D. J., Hodes, M. E., Norton, J. A., Connecally, P., & Christian, J. C. (1995). Cognitive scores in carriers of Huntington's disease gene compared to noncarriers. *AnnNeurol*, 37 (5), 657-64.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage*, 20, S8-S17.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition*, 27(3), 438–453. http://doi.org/10.3758/BF03211539

Friederici, A. D. (2017). Evolution of the neural language network. *Psychon Bull Rev* 24:41–47

Gagnon, M., Barrette, J., & Macoir, J. (2018). Language Disorders in Huntington Disease: A Systematic Literature Review. *Cognitive and Behavioral Neurology*, *31*(4), 179-192.

García, A.M., Sedeño, L., Trujillo, N., Bocanegra, Y., Gomez, D., Pineda, D., Villegas, A., Muñoz, E., Arias, W., & Ibáñez, A. (2017). Language deficits as a preclinical window into Parkinson's disease: evidence from asymptomatic parkin and dardarin mutation carriers. *Journal of the International Neuropsychological Society* 23, 150-158.

García, A.M., Bocanegra, Y., Herrera, E., Pino, M., Muñoz, E., Sedeño, L., & Ibáñez, A., (2018). Action-semantic and syntactic deficits in subjects at risk for Huntington's disease. *Journal of neuropsychology* (Mar 11, 2017). doi: 10.1111/jnp.12120.

Gayraud, F., H.-R. Lee, &Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics & Phonetics* 25(3): 198–209.

Golden, C. J., & Freshwater, S. M. (1978). *The Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. Chicago, IL: Stoelting.

Gomar, J.J., Ortiz-Gil, J., McKenna, P.J., Salvador, R., Sans-Sansa, B., Sarro, S., Guerrero, A., & Pomarol-Clotet, E. (2011). Validation of the Word Accentuation Test (TAP) as a means of estimating premorbid IQ in Spanish speakers. *Schizophrenia Research*, 128, 175-176.

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders*. Lea & Febiger.

Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston naming test*. Lea & Febiger.

Gordon, W. P., & Illes, J. (1987). Neurolinguistic characteristics of language production in Huntington's disease: a preliminary report. *Brain and Language*, 31(1), 1-10.

Graybiel, A. M. (1995). Building Action Repertoires: Memory and Learning Functions of the Basal Ganglia. *Current Opinion in Neurobiology*, 5, 733-741.

Hamilton, A., Ferm, U., Heemskerk, A. W., Twiston-Davies, R., Matheson, K. Y., Simpson, S. A., & Rae, D. (2012). Management of speech, language and communication difficulties in Huntington's disease. *Neurodegenerative Disease Management*, *2*(1), 67-77.

Hartelius, L., Jonsson, M., Rickeberg, A., & Laakso, K. (2010). Communication and Huntington's disease: qualitative interviews and focus groups with persons with Huntington's disease, family

members, and carers. *International journal of language & communication disorders*, 45(3), 381-393.

Hinzen, W., Rosselló, J., Morey, C., Cámara, E., García-Gorro, C., Salvador, R., & de Diego-Balaguer, R. (2018). A systematic linguistic profile of spontaneous narrative speech in pre-symptomatic and early stage Huntington's disease. *Cortex*, 100:71-83, doi: 10.1016/j.cortex.2017.07.022.

Ho, A.K., Sahakian, B.J., Brown, R.G., Barker, R.A., Hodges, J.R., Ané, M.N., Snowden, J., Thompson, J., Esmonde, T., Gentry, R., Moore, J.W., & Bodner, T. (2003). Profile of cognitive progression in early Huntington's disease. *Neurology*, 61(12), 1702- 1706.

Hodges, J. R., Salmon, D. P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: a controlled prospective study. *Journal of Neurology, Neurosurgery & Psychiatry*, *53*(12), 1089-1095.

Hodges, J. R., Salmon, D. P., & Butters, N. (1991). The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain*, *114*(4), 1547-1558.

Huntington Study Group (Kieburtz K, primary author). The Unified Huntington's Disease Rating Scale: Reliability and Consistency. *Movement Disorders*, 1996;11:136-142.

Illes, J. (1989). Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's, and Parkinson's. *Brain and Language*, 37(4), 628–642.

Jensen, A. M., Chenery, H. J., & Copland, D. A. (2006). A comparison of picture description abilities in individuals with vascular subcortical lesions and Huntington's disease. *Journal of communication disorders*, 39(1), 62-77.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149.

Klasner, E., & Yorkston, K. (2001). Linguistic and cognitive supplementation strategies as augmentative and alternative communication techniques in Huntington's disease: Case report. *Augmentative and Alternative Communication*, 17(3), 154-160.

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Science* 14(9), 392–399.

Lieberman, P. (2007). The evolution of human speech; Its Anatomical and neural bases. *Current Anthropology*, 48, 39-66.

Lipski, W. J., Alhourani, A., Pirnia, T., Jones, P. W., Dastolfo-Hromack, C., Helou, L. B., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Fiez, J. A., Turner, R. S., Richardson, R. M. (2017). Subthalamic nucleus neurons differentially encode early and late aspects of speech production. *Journal of Neuroscience*, *38*(24), 5620-5631.; doi: http://dx.doi.org/10.1101/227793.

Lombardo, M. V., Pierce, K., Eyler, L. T., CarterBarnes, C., Ahrens-Barbeau, C., Solso, S., Kampell, K., & Courchesne, E. (2015). Different functional neural substrates for good and poor language outcome in autism. *Neuron*, 86(2), 567–577.

Longworth, C. E., Keenan, S. E., Barker, R. A., Marslen-Wilson, W. D., & Tyler, L. K. (2005). The basal ganglia and rule-governed language use: evidence from vascular and degenerative conditions. *Brain*, 128(3), 584-596.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates. Web page: http://childes.psy.cmu.edu/clan

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., & Fazio, F. (2001). Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage*, 13(1), 110–118.

Murray, L. L., & Lenz, L. P. (2001). Productive syntax abilities in Huntington's and Parkinson's diseases. *Brain and Cognition*, 46(1), 213-219.

Nemeth, D., Dye, C.D., Sefcsik, T., Janacsek, K., Turi, Z., Londe, Z., Klivenyi, P., Kincses, Z.T., Szabó, N., & Vecsei, L. (2012). Language deficits in pre-symptomatic Huntington's disease: Evidence from Hungarian. *Brain and language* 121, 248-253.

Papoutsi, M., Labuschagne, I., Tabrizi, S. J., & Stout, J. C. (2014). The cognitive burden in Huntington's disease: pathology, phenotype, and mechanisms of compensation. *Movement Disorders*, 29(5), 673-683.

Paulsen, J.S., Long, J.D., Ross, C.A., Williams, J. K., Nance, M. A., Erwin, C.J., Westervelt, H.J., Harrington, D. L., Bockholt, H. J., Zhang, Y., McCusker, E. A., Chiu, E. M., Panegyres P. K., &the PREDICT-HD Investigators and Coordinators of the Huntington Study Group (2014). Prediction of manifest Huntington's disease with clinical and imaging measures: a prospective observational study. *Lancet Neurol*13(12):1193-201.

Paynter, J., & Peterson, C. (2010). Language and ToM development in autism versus Asperger syndrome: Contrasting influences of syntactic versus lexical/semantic maturity. *Research in Autism Spectrum Disorders*, 4(3), 377-385.

Podoll, K., Caspary, P., Lange, H. W., & Noth, J. (1988). Language functions in Huntington's disease. *Brain*,111(6), 1475-1503.

Poudel, G. R., Stout, J. C., Gray, M. A., Salmon, L., Churchyard, A., Chua, P., Borowsky, B., Egan, G. F.,& Georgiou-Karistianis, N. (2015). Functional changes during working memory in Huntington's disease: 30-month longitudinal data from the IMAGE-HD study. *Brain Structure and Function*, 220(1), 501-512.

Rosenstein, M., Foltz, P., DeLisi, L., & Elvevåg, B. (2015). Language as a biomarker in those at high-risk for psychosis. *Schizophrenia Research*, 165, 249-250.

Rusz, J., Saft, C., Schlegel, U., Hoffman, R., & Skodda, S. (2014). Phonatory dysfunction as a preclinical symptom of Huntington disease. *PloS one*, 9(11), e113412.

Saft, C., Lissek, S., Hoffmann, R., Nicolas, V., Tegenthoff, M., Juckel, G., & Brüne, M. (2013). Mentalizing in preclinical Huntington's disease: an fMRI study using cartoon picture stories. *Brain imaging and behavior*, 7(2), 154-162.

Saldert, C., Eriksson, E., Petersson, K., &Hartelius, L. (2010). Interaction in conversation in Huntington´ s disease: An activity-based analysis and the conversation partner's view of change. *Journal of Interactional Research in Communication Disorders*, *1*(2), 169-197.

Sambin, S., Teichmann, M., de Diego Balaguer, R., Giavazzi, M., Sportiche, D., Schlenker, P., & Bachoud-Levi, A. C. (2012). The role of the striatum in sentence processing: disentangling syntax from working memory in Huntington's disease. *Neuropsychologia*, 50(11), 2625-2635.

Silverman, G. (1973). Redundancy, repetition and pausing in schizophrenic speech. *The British Journal of Psychiatry,* 122(569), 407-413.

Simonyan, K., & Fuertinger, S. (2015). Speech networks at rest and in action: interactions between functional brain networks controlling speech production. *Journal of Neurophysiology*, *113*(7), 2967–2978.

Simonyan, K., Ackermann, H., Chang, E. F., & Greenlee, J. D. (2016). New Developments in Understanding the Complexity of Human Speech Production. *The Journal of Neuroscience*, *36*(45), 11440–11448.

Skodda, S., Grönheit, W., Lukas, C., Bellenberg, B., von Hein, S. M., Hoffmann, R., & Saft, C. (2016). Two different phenomena in basic motor speech performance in premanifest Huntington disease. *Neurology,* 86(14), 1329-1335.

Solomon, A. C., Stout, J. C., Johnson, S. A., Langbehn, D. R., Aylward, E. H., Brandt, J., Ross, C. A., Beglinger, L., Hayden, M.R., Kieburtz, K., Kayson, E., Julian-Baros, E., Duff, K., Guttman, M., Nance, M., Oakes D., Shoulson, I., Penziner, E., Paulsen, J. S, &the PREDICT-HD Investigators and Coordinators of the Huntington Study Group. (2007). Verbal Episodic Memory Declines Prior to Diagnosis in Huntington's Disease. *Neuropsychologia*, *45*(8),1767–1776.

Steele, S., Joseph, R. M., & Tager-Flusberg, H. (2003). Brief report: Developmental change in theory of mind abilities in children with autism. *Journal of Autism and Developmental Disorders*, 33(4), 461-467.

Stout, J. C., Paulsen, J. S., Queller, S., Solomon, A. C., Whitlock, K. B., Campbell, J. C., Carlozzi, N. D., Kevin Beglinger, L.J., Langbehn, D. R., Biglan, Kevin, M., & Aylward, E. H. (2011). Neurocognitive Signs in Prodromal Huntington Disease. *Neuropsychology*, 25(1), 1-14.

Swinney, D., Zurif, E., Prather, P., & Love, T. (1996). Neurological distribution of processing resources underlying language comprehension. *Journal of Cognitive Neuroscience*, *8*(2), 174-184.

Tabrizi, S.J., Scahill, R.I.,& Owen, G, et al. (2013). Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. *Lancet Neurol*12:637-649.

Teichmann, M., Dupoux, E., Cesaro, P., &Bachoud-Lévi, A. C. (2008). The role of the striatum in sentence processing: evidence from a priming study in early stages of Huntington's disease. *Neuropsychologia*, 46(1), 174-185.

Teichmann, M., Dupoux, E., Kouider, S., & Bachoud-Lévi, A. C. (2006). The role of the striatum in processing language rules: evidence from word perception in Huntington's disease. *Journal of Cognitive Neuroscience*, 18(9), 1555-1569.

Teichmann, M., Dupoux, E., Kouider, S., Brugières, P., Boissé, M. F., Baudic, S., Cesaro, P., Peschanski, M., & Bachoud-Lévi, A. C. (2005). The role of the striatum in rule application: The model of Huntington's disease at early stage. *Brain*, 128(5), 1155–1167. http://doi.org/10.1093/brain/awh472

Teichmann, M., Gaura, V., Démonet, J. F., Supiot, F., Delliaux, M., Verny, C., Renou, P. & Bachoud-Lévi, A. C. (2008). Language processing within the striatum: evidence from a PET correlation study in Huntington's disease. *Brain*, 131(4), 1046-1056.

Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of clinical neuropsychology*, 19(2), 203-214.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9(2),266–276.

Vogel, A. P., Shirbin, C., Churchyard, A. J., & Stout, J. C. (2012). Speech acoustic markers of early stage and prodromal Huntington's disease: A marker of disease onset? *Neuropsychologia*, 50(14), 3273–3278. http://doi.org/10.1016/j.neuropsychologia.2012.09.011.

Walker, M. A. (1996). Limited attention and discourse structure. *Comput. Linguist.* 22, 255–264.

Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised*. Psychological Corporation.

Wiecki, T.V., Antoniades, C.A., Stevenson, A., Kennard, C., Borowsky, B., Owen, G., Leavitt, B., Ross, R., Durr, A., Trabizi, S., & Frank, M. (2016). A Computational Cognitive Biomarker for Early-Stage Huntington's Disease. *PLoS ONE*, 11(2): e0148409. doi:10.1371/journal.pone.0148409.

Wiltschko, M. (2018) Discovering syntactic variation. In: Hornstein, N., Lasnik, H., Patel-Grosz, P., Yang, Ch. (eds.), Syntactic *Structures after 60 Years. The Impact of the Chomskyan Revolution in Linguistics.* Series: Studies in Generative Grammar 129. pp 427-460.

Wright, H. H., Downey, R. A., Gravier, M., Love, T., & Shapiro, L. P. (2007). Processing distinct linguistic information types in working memory in aphasia. *Aphasiology*, *21*(6-8), 802-813.