Arts & Sciences Electronic Theses and Dissertations

Arts & Sciences

Spring 5-15-2015

The Attentional Control of Reading: Insights from Behavior, Imaging and Development

Sarah Ihnen
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Biology Commons

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology & Biomedical Sciences

Neurosciences

Dissertation Examination Committee:
Bradley Schlaggar, Chairperson
David Balota
Deanna Barch
Tamara Hershey
Steven Petersen
Rebecca Treiman

The Attentional Control of Reading:
Insights from Behavior, Imaging and Development

by

Sarah Katherine Zeller Ihnen

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2015

St. Louis, Missouri

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Our achievements, indeed our lives, result from the strange and capricious alchemy of circumstance, people and whatever meager virtue lies within us. Of my own circumstances and virtue, I have little to say. But there are many people in my life to whom I owe a debt of gratitude, and I would be remiss not to thank them at this moment of culmination.

I first felt the pull of Science in the ninth grade classroom of Maria Rosa Arbona, who showed me that the better part of biology is art, and whose love for her craft reinforced for me that I better figure out what it is I really want to do, and do it well. Later, the Wells Scholar Program at Indiana University afforded me with four of the most thrilling years of my life. Being a Wells Scholar was an opportunity to develop a broad and deep liberal arts education surrounded by the highest caliber of peers, all the while living in a sprawling, gorgeous and lively campus grounded in Midwestern friendliness. A little piece of me will always reside in the hills of southern Indiana, the place where, in addition to attaining a liberal arts education, I met my very best friends, fell in love with cycling and found the love of my life.

I am grateful to Dr. Greg DeAngelis, for advocating for my acceptance into the Neuroscience PhD program here at Washington University, for keeping in contact with me during my deferral, and for encouraging me to find my place in the program, and in the Cognitive, Computational and Systems Neuroscience (CCSN) pathway. I am not sure that I would be here today were it not for his initial, perhaps unwarranted, belief that I could succeed in graduate school.

During my time in the Neuroscience program, I have been supported by many funding sources. I am grateful for all of that support, including the CCSN pathway National Science Foundation Integrative Graduate Education and Research Trainee Fellowship (0548890). I also want to acknowledge the National Instituted of Health R01 training grant awarded to Brad (HD057076) for the many, many scanning sessions that it funded.

When I first began my studies at Wash U, I rotated through several excellent laboratories. In Dr. Maurizio Corbetta's lab, I first realized I wanted to be a physician scientist, particularly after interacting with Chad Sylvester and Guarav Patel (now, both MD/PhD). Rotating with Dr. Larry Snyder provided me with useful exposure to the technique of single-unit recordings in monkeys, and Larry impressed me with his thoughtful engagement with his students. I seriously considered working with Dr. Randy Buckner, before he was recruited away from Washington University, and I am grateful that I had the opportunity to learn from him. Randy was an enthusiastic, hard-working mentor to me during my rotation, and also very encouraging of my decision to apply to the MD/PhD program.

Choosing the climate in which to pursue one's graduate training – including both the mentor and the lab dynamic – is certainly the paramount determinant of the productivity (and enjoyment) of an extended and formative period of career development. Room 2220 in the East building is a most auspicious place in which to train as a graduate student or postdoc. Just as 2 popes are sometimes better than 1 (or not)[1], the Petersen/Schlaggar laboratory benefits from the joint and benevolent ruling of Steve and Brad, and together they foster a climate where people are provided with abundant resources, supported in their intellectual wanderings, and continuously pushed to be rigorous and thoughtful with their science. Theirs is a lab with a Social Chairperson, a well-used lunch table, and no dearth of vigorous and frequent rhetorical exchanges on every imaginable topic, neuroscientific and otherwise. In short, lab has been a fun and fruitful place for me to be during these past few years, and I will miss being there.

My asking Brad to be his student was the fortuitous consequence of Randy Buckner having left Wash U at the time I was deciding to which lab I'd commit. I first met "Dr. Schlaggar" during my qualifying exam, and I was struck by his combination of shrewd intellect and disarming approach to the exchange of ideas. I had a hunch that Brad's likely mentoring style

---

[1] C.f. the dual papacy (or Western Schism) of 1378-1417

would mesh with my strengths and weaknesses, and my bet has certainly paid off. Brad is the consummate physician scientist, in that he is able to juggle, deftly, many competing demands at once, and has the work ethic and stamina to direct thoughtful attention to each demand. More impressive than Brad's juggling skills, however, is his deep integrity – with regard both to pursuing a scientific question and, by all accounts, to taking care of patients. I've never questioned Brad's dedication to helping me succeed, whether it's evidenced by his 0600 email replies to my queries or by his willingness – eagerness, actually – to make time for me to discuss an analysis result or hash out an interpretation. And while I would never go so far as to accuse Brad of religiosity (cue the pope joke on the previous page), it is my opinion that he lives and breathes the ethos of service that was such a part of my childhood Catholic education, in that he works tirelessly, doing something at which he is very talented, because he believes in the importance of making a contribution with one's life. I totally admire and appreciate Brad's devotion to his work, and I count myself lucky to have been one of its beneficiaries.

I've also been especially fortunate to benefit from having a second primary mentor all these years, Steve Petersen. While Steve and I may show little overlap in the assorted Venn diagrams of our personality traits, hobbies and food preferences (see Supplemental Figure 1 – just joking), I would like to think that I have made responsible use of the privilege to learn from him, as he is truly a scientist of stunning acumen and indefatigable fidelity to the truth. Over the development of this project, he has pushed me to think critically, in ways that may have been convenient to avoid, about my data. I have always been especially appreciative of the fact that he made time for me on a regular basis, despite my not being "his student," officially. I regret that by leaving the lab, I will be depriving Steve of one of his main dialectic targets, my bicycle. In all seriousness, although I did not always thoroughly enjoy each argument with Steve while it was occurring, I am ever so grateful for having had the opportunity over the last several years to

hone my ability to defend my position, scientifically or otherwise, against a theoretically intimidating opponent.

The other four members of my thesis committee have also shared liberally and kindly with me their perspicacity and stockpiles of expertise. I am grateful to Dr. Dave Balota, Dr. Becky Thiemann, Dr. Deanna Barch and Dr. Tammy Hershey for sitting through eight (but who counted?) thesis committee update meetings with me, and I appreciate that each of them has seemed consistently genuinely interested in my project and in its eventual completion. Dave and Becky were particularly closely involved with the erstwhile manuscript associated with Chapter 3, and although our back-and-forth was sometimes painful in what it required of me, I know that they've helped make me a better scientist, and I am thankful for their careful reading and thinking about my work. Deanna and Tammy were helpful in shaping the imaging analyses described in Chapter 2, and they both also took care to keep me focused and on track throughout the last few years. I benefitted tremendously from the diversity of perspectives and knowledge bases of all 6 of my thesis committee members.

I am so very grateful to have been supported during my time in the MD/PhD program by the truly talented members of the Wash U MSTP administrative staff – Brian Sullivan, Christy Durbin, Liz Bayer and Linda Perniciario (and previously Andrew Richards). These individuals combine eagle-eyed attention to detail and parental-like protectiveness in such an unobtrusively effective manner that it would be easy to take for granted how much their assistance has facilitated my progress along this challenging dual degree pathway. I wish to take the time to thank them all now, both for their quiet hard work and for their kindness. Similarly, I am appreciative of the encouragement and support of Anna Cook, whom I first met in Randy Buckner's lab many years ago, and who has continued to lend both a listening ear and words of advice at various points over these past 10 years.

I am so pleased to have crossed paths with Jessica Church-Lang. She picked me up from Lambert Airport when I first visited St. Louis to interview for the PhD program. I am grateful to her for teaching me, at the beginning of graduate school, a great deal about how things work in the lab, scientifically and otherwise. She continues to inspire me now, from a distance, by her unassuming industriousness and clarity of thinking. I would also like to thank her for the thoughtful feedback that she provided on Chapter 2 of this thesis.

I am also fortunate to have followed in the footsteps of Alecia Vogel-Hammen, who is a role model to me as she launches into her clinical training to become a child psychiatrist. Her enthusiasm and positive attitude are infectious, and her sharp mind and diligence have served as a prime example, to me, of what I'd like to be "when I grow up." We have shared many spirited discussions on the topics of lexical processing and development, and I am grateful for her thoughtful and incisive feedback on my work, especially regarding the contents of this thesis.

I would like to thank all the other past members of the Petersen and Schlaggar labs, including Damian Fair, Steve Nelson, Alex Cohen and Kelly Barnes. Steve Nelson, in particular, has been someone whose attention to detail and effusive passion for neuroscience has served as an inspiration to me.

My time in lab has also been made more enjoyable and scientifically richer because of the intellectual curiosity (and unmatched lunchtime companionship) of the current Petersen and Schlaggar lab members, including Tunde Adeyemo, Stephanie Berk, Joe Dubis, Evan Gordon, Deanna Greene, Tim Laumann, Binyam Nardos, Maital Neta, Jonathan Power and Haoxin Sun. I will very much miss the regular company of these folks, even though I imagine that we haven't seen the last of each other.

I am also grateful to Mary Downey-Jones, Fran Miezin, Becky Coalson, Kelly McVey and Becca Lepore for their kind help during my years in lab. Mary never made me feel like I was asking a silly question, and she kept my study participants paid and happy. Fran provided a

tremendous amount of technical support, although I am still not sure how it is that he frequently managed to "fix" my computer simply by standing within 2 feet of it ("the Fran effect"). Becky was the person in lab who so patiently introduced me to fIDL and general lab procedure back in 2005, and she has continued to help me in numerous ways since then, including endless (and complaint-free) updating of the lab EndNote library. Kelly provided invaluable assistance as well, including the hand-scoring of thousands of reaction times; her thoroughness and reliability with that task was something I never had to second-guess. Becca, along with Becky and Kelly, graciously and skillfully assisted me with data collection for over 90 MRI study subjects, at times also pinch-hitting for me with the cognitive testing, as when I somehow had to work through 5 subjects sessions in a row.

I would also like to thank Hristina Lessov-Schlaggar for several acts of generosity. First, she has unfailingly tolerated my occasional intrusions into Schlaggar family time, at times even graciously welcoming me into the Schlaggar home for my after-hours queries and concerns. Hristina has also kindly assisted me with interpreting some of the more complicated aspects of the statistical analyses in this thesis. She is a role model to me in other ways as well, including the graceful manner in which she juggles the demands of her own career, combined with the challenges of being married to a very busy person, all the while raising two lovely and accomplished children.

Speaking of children, I do want to express my appreciation for all the children and adults who participated in my studies. I interacted with at least 170 study participants to acquire the data for this thesis, and I am glad that all of these volunteers (along with the parents of the children) decided that it was interesting and useful to participate in the scientific enterprise by giving of their time and effort.

The friends and family whose love and support I would like to recognize are too numerous to mention here in full, but I'd like to highlight a few. I mentioned earlier those very

best friends from Bloomington. If it's true that you can tell a lot about a person by the company she keeps, then I don't ever want to be far from Sara Moellers Kleinschmidt, Anna Barry Cope, Crystal Spivey Wildeman and Tracey Guildenbecher Columb. Ten years out, I still count these four women as the very best "college friends" a person could want. I rely on them for emotional sustenance, an unbecoming amount of laughter when we are together and, importantly, both implicit and explicit support of my decision to pursue a demanding career.

The people with whom I started the Wash U MSTP in 2006 have also been a source of encouragement during my thesis work. In particular, it has been rejuvenating (not to mention, fun) to have regular get-togethers with the other women my cohort, and I'm so thankful especially for the friendship and camaraderie of Susan Canny, Aggie Bielska, Kristin Bibee and Monqiue Bruinsma.

After Lucy was born in August 2008, I've only been able to come to work because she was at home (eventually along with her brother, Oscar) in the hands (literally, often) of the superstars who have been our nannies. I can't express my gratitude heartily enough for the relationships each of these women built with my children, and for the love and trust they earned from me and Alex. I will long be indebted to Amanda Felber, Carissa Bartholomew, Jamie Neidorf, Christine Gould, and Caitlyn McEvilly. Caitlyn, our current nanny, has truly been Mary Poppins for us during these last few months of thesis preparation, when I've worried about being away from Lucy and Oscar for extended periods of time. Caitlyn has made it much, much easier not to worry.

I come from a large and generally not dysfunctional extended family as well, and I am especially grateful to Aunt Becky, Aunt Carol and Uncle Tom, Meighan, Hilary, Hannah and Olivia for a lifetime of support, encouragement and family parties that made the trip home worthwhile. Years of their cheerleading from the sidelines has not gone unnoticed.

We don't get to choose the people with whom we share our parents, but if we could, I'd choose each of my siblings, Ben, Becky, Micah and Sam. I thank them for allowing me to cajole them into playing board games all these years. I thank them for alternately putting me in my place and for being my biggest fans, as needed. It has been especially wonderful to have had Micah living in St. Louis these last 5 years, as there is something about having your sibling sleep just a couple of miles away that simply feels like the way things should be. I look forward to continue sharing in all of their lives.

How do I possibly begin to express my gratitude for my mom and dad? I could thank them for filling my childhood with books[2] and kindness and ideas and music and an abiding sense of security. I could thank them for being my relentless champions, extraordinarily and selflessly, or for letting me live in Italy with my boyfriend (now husband) after graduating from college, so that I could ride my bicycle, learn how to cook and study Italian – I'm not sure they know what a gift that was. I could thank them for showing me how to *lean in*, before that phrase became a book title and a buzzword, and for teaching me, by example, that you shouldn't do something or agree with something just because it's popular. I could thank them for being people of substance and compassion, or for planting a seed somewhere inside me that makes me badly want to be a person of substance and compassion. Try as I may, I will never be able to thank them enough for all of that, or to pay them back. But if, by the end of my days, I've made my children and my husband feel half as loved as they've made me feel, I'll have done one big thing, *the* big thing, right.

My husband, Alex, is my past, my present and my future. He loves me for who I am, all the while making me want to be a better person, which is perhaps the most we can truly ask of

---

[2] Yes, a second footnote, and we're still in the Acknowledgments. Please excuse my parenthetical proclivities for the next couple of hundred pages. I vividly recall, as a child, pulling the Scientific American volume *Images of Mind*, by Mike Posner and Marc Raichle, off the shelf in my parents' library. I never would have guessed that sometime down the road, I would meet both of those scientists, and have the chance to interact with so many other brilliant and thoughtful people, present readership included.

another person. I thank him for his steadfast support of my goals, in ways big and small, and for being my courage when needed. For seven years and counting, he has made significant personal and professional sacrifices to provide for our family, a devotion I don't take for granted. Eventually, I hope to be able to return the favor. Together, we will never just sit around waiting to die, and that's one of the many reasons I love him.

DEDICATION

Lucy and Oscar, you have infused into the deepest part of me, into a space I didn't know existed, an awareness of being alive that is as vast as it is acute, as visceral as it is transcendent, and vulnerable as it is fierce. Notwithstanding the thousands of generations of humanity that preceded us, I cannot convince myself that anyone else has ever felt about another person the way I feel about you two. This work is in your honor.

ABSTRACT OF THE DISSERTATION

The Attentional Control of Reading:
Insights from Behavior, Imaging and Development

by

Sarah Katherine Zeller Ihnen

Doctor of Philosophy in Biology & Biomedical Sciences
(Neurosciences)

Washington University in St. Louis, 2015

Professor Bradley L. Schlaggar, Chairperson

The process by which the initially attention-requiring task of transforming scribbles into meaningful concepts eventually becomes facile remains a central riddle of cognitive neuroscience. This body of work represents an effort to provide forward movement in answering the question of how attentional control mediates the process of reading, both by considering different stages of reading competence (development) and by seeking convergence between types of evidence (behavior and imaging).

Inspired by a study published by Balota and colleagues in 2000, the paradigm used throughout this work involves comparing a simple speeded reading task vs. a regularize ("sound out") task (Balota et al. 2000). In the first data chapter, I replicate the essential findings of the Balota et al. study in 2 young adult cohorts, confirming that stimulus characteristics, including lexicality and frequency, influence reading task performance in a manner that is modulated by top-down attentional control. I furthermore argue that the reaction time (RT) patterns are consistent with 2 distinct mechanisms by which top-down attentional control interacts with reading processes, pathway control and response checking. I then present evidence, motivated

by the 2-mechanism hypothesis, that 2 sets of brain regions, including members of previously defined attentional control networks, show separable activity patterns that map nicely onto roles reflecting pathway control and response checking.

In the second data chapter, I show that 8-10 year old children, like young adults, can perform the regularize task. Unexpectedly, the early readers are faster than the experienced readers to regularize, and this speed advantage for children holds for both words and pseudowords. Because children are slower than adults across a range of cognitive tasks (e.g., Kail 1991) – with children showing particular immaturity with regard to inhibiting prepotent responses (e.g., Davidson et al. 2006) – the developmental observation is remarkable in and of itself. Complemented by a cadre of *post hoc* analyses, the age groups differences can also be interpreted as additional support for the 2-mechanism interaction of attention and reading.

Together, these results suggest that dissociable subcomponents of attentional control interact with subcomponents of reading processing, and that these interactions are dynamic across skill development and across task demands.

EPIGRAPH

     As I was choosing the word stimuli for these studies, I discovered this delightful poem. Trenité's verses artfully illustrate that the English language is rife with spelling-to-sound irregularities, a fact that lends both beauty and madness to the acts of reading and spelling. On occasion, madness breeds insight, a possibility I explore in this dissertation by leveraging the madness (spelling-to-sound irregularity) to illuminate the beauty (reading).

The Chaos, by Gerald Nolst Trenité (1922)

Dearest creature in creation
Studying English pronunciation,
I will teach you in my verse
Sounds like corpse, corps, horse and worse.

I will keep you, Susy, busy,
Make your head with heat grow dizzy;
Tear in eye, your dress you'll tear;
Queer, fair seer, hear my prayer.

Pray, console your loving poet,
Make my coat look new, dear, sew it!
Just compare heart, hear and heard,
Dies and diet, lord and word.

Sword and sward, retain and Britain
(Mind the latter how it's written).
Made has not the sound of bade,
Say-said, pay-paid, laid but plaid.

Now I surely will not plague you
With such words as vague and ague,
But be careful how you speak,
Say: gush, bush, steak, streak, break, bleak.

Previous, precious, fuchsia, via
Recipe, pipe, studding-sail, choir;
Woven, oven, how and low,
Script, receipt, shoe, poem, toe.

Say, expecting fraud and trickery:
Daughter, laughter and Terpsichore,
Branch, ranch,measles, topsails, aisles,
Missiles, similes, reviles.

Wholly, holly, signal, signing,
Same, examining, but mining,
Scholar, vicar, and cigar,
Solar, mica, war and far.

From "desire": desirable-admirable from "admire",
Lumber, plumber, bier, but brier,
Topsham, brougham, renown, but known,
Knowledge, done, lone, gone, none, tone,

One, anemone, Balmoral,
Kitchen, lichen, laundry, laurel.
Gertrude, German, wind and wind,
Beau, kind, kindred, queue, mankind.

Tortoise, turquoise, chamois-leather,
Reading, Reading, heathen, heather.
This phonetic labyrinth
Gives moss, gross, brook, brooch, ninth, plinth.

Have you ever yet endeavoured
To pronounce revered and severed,
Demon, lemon, ghoul, foul, soul,
Peter, petrol and patrol?

Billet does not end like ballet;
Bouquet, wallet, mallet, chalet.
Blood and flood are not like food,
Nor is mould like should and would.

Banquet is not nearly parquet,
Which exactly rhymes with khaki.
Discount, viscount, load and broad,
Toward, to forward, to reward,

Ricocheted and crocheting, croquet?
Right! Your pronunciation's OK.
Rounded, wounded, grieve and sieve,
Friend and fiend, alive and live.

Is your r correct in higher?
Keats asserts it rhymes Thalia.
Hugh, but hug, and hood, but hoot,
Buoyant, minute, but minute.

Say abscission with precision,
Now: position and transition;
Would it tally with my rhyme
If I mentioned paradigm?

Twopence, threepence, tease are easy,
But cease, crease, grease and greasy?
Cornice, nice, valise, revise,
Rabies, but lullabies.

Of such puzzling words as nauseous,
Rhyming well with cautious, tortious,
You'll envelop lists, I hope,
In a linen envelope.

Would you like some more? You'll have it!
Affidavit, David, davit.
To abjure, to perjure. Sheik
Does not sound like Czech but ache.

Liberty, library, heave and heaven,
Rachel, loch, moustache, eleven.
We say hallowed, but allowed,
People, leopard, towed but vowed.

Mark the difference, moreover,
Between mover, plover, Dover.
Leeches, breeches, wise, precise,
Chalice, but police and lice,

Camel, constable, unstable,
Principle, disciple, label.
Petal, penal, and canal,
Wait, surmise, plait, promise, pal,

Suit, suite, ruin. Circuit, conduit
Rhyme with "shirk it" and "beyond it",
But it is not hard to tell
Why it's pall, mall, but Pall Mall.

Muscle, muscular, gaol, iron,
Timber, climber, bullion, lion,
Worm and storm, chaise, chaos, chair,
Senator, spectator, mayor,

Ivy, privy, famous; clamour
Has the a of drachm and hammer.
Pussy, hussy and possess,
Desert, but desert, address.

Golf, wolf, countenance, lieutenants
Hoist in lieu of flags left pennants.
Courier, courtier, tomb, bomb, comb,
Cow, but Cowper, some and home.

"Solder, soldier! Blood is thicker",
Quoth he, "than liqueur or liquor",
Making, it is sad but true,
In bravado, much ado.

Stranger does not rhyme with anger,
Neither does devour with clangour.
Pilot, pivot, gaunt, but aunt,
Font, front, wont, want, grand and grant.

Arsenic, specific, scenic,
Relic, rhetoric, hygienic.
Gooseberry, goose, and close, but close,
Paradise, rise, rose, and dose.

Say inveigh, neigh, but inveigle,
Make the latter rhyme with eagle.
Mind! Meandering but mean,
Valentine and magazine.

And I bet you, dear, a penny,
You say mani-(fold) like many,
Which is wrong. Say rapier, pier,
Tier (one who ties), but tier.

Arch, archangel; pray, does erring
Rhyme with herring or with stirring?
Prison, bison, treasure trove,
Treason, hover, cover, cove,

Perseverance, severance. Ribald
Rhymes (but piebald doesn't) with nibbled.
Phaeton, paean, gnat, ghat, gnaw,
Lien, psychic, shone, bone, pshaw.

Don't be down, my own, but rough it,
And distinguish buffet, buffet;
Brood, stood, roof, rook, school, wool, boon,
Worcester, Boleyn, to impugn.

Say in sounds correct and sterling
Hearse, hear, hearken, year and yearling.
Evil, devil, mezzotint,
Mind the z! (A gentle hint.)

Now you need not pay attention
To such sounds as I don't mention,
Sounds like pores, pause, pours and paws,
Rhyming with the pronoun yours;

Nor are proper names included,
Though I often heard, as you did,
Funny rhymes to unicorn,
Yes, you know them, Vaughan and Strachan.

No, my maiden, coy and comely,
I don't want to speak of Cholmondeley.
No. Yet Froude compared with proud
Is no better than McLeod.

But mind trivial and vial,
Tripod, menial, denial,
Troll and trolley, realm and ream,
Schedule, mischief, schism, and scheme.

Argil, gill, Argyll, gill. Surely
May be made to rhyme with Raleigh,
But you're not supposed to say
Piquet rhymes with sobriquet.

Had this invalid invalid
Worthless documents? How pallid,
How uncouth he, couchant, looked,
When for Portsmouth I had booked!

Zeus, Thebes, Thales, Aphrodite,
Paramour, enamoured, flighty,
Episodes, antipodes,
Acquiesce, and obsequies.

Please don't monkey with the geyser,
Don't peel 'taters with my razor,
Rather say in accents pure:
Nature, stature and mature.

Pious, impious, limb, climb, glumly,
Worsted, worsted, crumbly, dumbly,
Conquer, conquest, vase, phase, fan,
Wan, sedan and artisan.

The th will surely trouble you
More than r, ch or w.
Say then these phonetic gems:
Thomas, thyme, Theresa, Thames.

Thompson, Chatham, Waltham, Streatham,
There are more but I forget 'em-
Wait! I've got it: Anthony,
Lighten your anxiety.

The archaic word albeit
Does not rhyme with eight-you see it;
With and forthwith, one has voice,
One has not, you make your choice.

Shoes, goes, does *. Now first say: finger;
Then say: singer, ginger, linger.
Real, zeal, mauve, gauze and gauge,
Marriage, foliage, mirage, age,

Hero, heron, query, very,
Parry, tarry fury, bury,
Dost, lost, post, and doth, cloth, loth,
Job, Job, blossom, bosom, oath.

Faugh, oppugnant, keen oppugners,
Bowing, bowing, banjo-tuners
Holm you know, but noes, canoes,
Puisne, truism, use, to use?

Though the difference seems little,
We say actual, but victual,
Seat, sweat, chaste, caste, Leigh, eight, height,
Put, nut, granite, and unite.

Reefer does not rhyme with deafer,
Feoffer does, and zephyr, heifer.
Dull, bull, Geoffrey, George, ate, late,
Hint, pint, senate, but sedate.

Gaelic, Arabic, pacific,
Science, conscience, scientific;
Tour, but our, dour, succour, four,
Gas, alas, and Arkansas.

Say manoeuvre, yacht and vomit,
Next omit, which differs from it
Bona fide, alibi
Gyrate, dowry and awry.

Sea, idea, guinea, area,
Psalm, Maria, but malaria.
Youth, south, southern, cleanse and clean,
Doctrine, turpentine, marine.

Compare alien with Italian,
Dandelion with battalion,
Rally with ally; yea, ye,
Eye, I, ay, aye, whey, key, quay!

Say aver, but ever, fever,
Neither, leisure, skein, receiver.
Never guess-it is not safe,
We say calves, valves, half, but Ralf.

Starry, granary, canary,
Crevice, but device, and eyrie,
Face, but preface, then grimace,
Phlegm, phlegmatic, ass, glass, bass.

Bass, large, target, gin, give, verging,
Ought, oust, joust, and scour, but scourging;
Ear, but earn; and ere and tear
Do not rhyme with here but heir.

Mind the o of off and often
Which may be pronounced as orphan,
With the sound of saw and sauce;
Also soft, lost, cloth and cross.

Pudding, puddle, putting. Putting?
Yes: at golf it rhymes with shutting.
Respite, spite, consent, resent.
Liable, but Parliament.

Seven is right, but so is even,
Hyphen, roughen, nephew, Stephen,
Monkey, donkey, clerk and jerk,
Asp, grasp, wasp, demesne, cork, work.

A of valour, vapid vapour,
S of news (compare newspaper),
G of gibbet, gibbon, gist,
I of antichrist and grist,

Differ like diverse and divers,
Rivers, strivers, shivers, fivers.
Once, but nonce, toll, doll, but roll,
Polish, Polish, poll and poll.

Pronunciation-think of Psyche!-
Is a paling, stout and spiky.
Won't it make you lose your wits
Writing groats and saying "grits"?

It's a dark abyss or tunnel
Strewn with stones like rowlock, gunwale,
Islington, and Isle of Wight,
Housewife, verdict and indict.

Don't you think so, reader, rather,
Saying lather, bather, father?
Finally, which rhymes with enough,
Though, through, bough, cough, hough, sough, tough??

Hiccough has the sound of sup...
My advice is: GIVE IT UP!

CHAPTER 1: INTRODUCTION AND BACKGROUND

In this chapter, several of the essential themes of this thesis will be introduced. In the first part, I will discuss topics related to the relevance of reading research, as well as the tools, models and priors from which this research emerged. First, I will briefly argue for the general importance of reading research, based on the observation that literacy is an important predictor of educational and economic success (Baer et al. 2009), coupled with my own opinion that reading is an intrinsically interesting cognitive phenomenon. I will then highlight the fact that impaired reading is a substantial public health concern, further underscoring the need for research on typical reading development. Next, I will delve into some of the issues related to mental chronometry, the use of reaction time measurements to investigate cognition, particularly as the technique pertains to investigations of lexical processing. I will then highlight a few of the more consistent findings with regard to the functional neuroanatomy of reading although, as will be revealed, Chapter 2 underscores the contributions of different sets of brain regions to reading tasks from those that have typically been emphasized. Turning from the brain to computational models, I will briefly describe the important features of dual route and connectionist architectures, emphasizing the tremendous utility of these theoretical frameworks. I will then speculate that computational models have contributed to the perpetuation of an automaticity dogma in the field of reading, perhaps to the detriment of novel conceptual advances. I will also make an argument that along with the automaticity dogma, there exists something of a phonology dogma in the literature. I will furthermore address the importance of choosing and accurately describing "reading-related" tasks, and justify the choice to employ a speeded naming task and a regularize task in the present work.

The second part of this chapter will elaborate on a few of the issues relevant to the developmental component of this work. First, I will posit one interesting way in which a line of investigation of dyslexia may provoke a conceptual approach to studying normal reading, a

reverse of the typical situation in which understanding the healthy system informs the disordered system. A clarification about the difference between development and skill acquisition will then be offered, followed by a brief discussion of the observation that development entails both general decreases in the time required to respond to stimuli and specific improvements in the various sub-components of what is referred to as executive, cognitive or attentional control. As will be discussed in greater detail in Chapter 3, the admixture of sometimes counter-modulatory features of the development of reading processing (e.g., consolidation of lexical knowledge) with features of the development of attentional control (e.g., improvements in inhibitory control) complicates the predictions and subsequent evaluations of age group differences in the present data.

Finally, I will conclude this chapter by outlining a road map for the 2 data chapters and by presaging the primary observations and conclusions of those chapters.

Reading research: relevance, tools, models and priors

The first section of this chapter surveys a few of the issues related to reading research in a way that sets the stage for the work presented in Chapters 2 and 3.

Reading research is important and fascinating science

Reading, the act of obtaining meanings from written language, is fundamental both for the transmission and preservation of the culture of our species and, for humans as individuals, for the task of navigating in an increasingly complex and abstract world. The riddle of how we read remains unsolved despite a long history of inquiry on the part of psycholinguists, cognitive neuroscientists and neurologists (e.g., Cattell 1886; Dejerine 1891, 1892; Pugh et al. 1996), marked even today by vigorous debate over questions as fundamental as whether or not particular pieces of cortex are dedicated to performing reading processing (contrast Cohen et al.

2002; Glezer et al. 2009 vs. Price and Devlin 2003; Price and Devlin 2011; Vogel et al. 2012)

Reading is of intrinsic interest in part because it represents a complex, learned cognitive skill that evolved only in the last few thousand years, presumably co-opting neural hardware adapted to support other functions. Furthermore, the substrate of reading – the written word – is manipulable in concrete and describable ways, such that reading offers a sort of model system for interrogating more general cognitive neuroscience phenomena, such as perceptual expertise and behavioral variability (across tasks, across the lifespan, across populations, etc.).

For example, it has been proposed that the protracted development of brain regions supporting cognitive control (e.g., Huttenlocher and Dabholkar 1997; Casey et al. 2000; Gogtay et al. 2004; Giedd 2004) confers a benefit to spoken language learning (Chrysikou et al. 2011), in that bottom-up, data-driven exploration of the environment by children – unconstrained by the biases of the top-down, rule-based thinking that characterizes cognitive maturity – affords an adaptive flexibility in extracting the statistical rules governing syntax, for example. This intriguing hypothesis, forwarded by Thompson-Schill and colleagues, provokes the speculation that observations regarding the development of reading, specifically, may illuminate the broader question of whether perceptual expertise is indeed a Faustian bargain, whereby certain component abilities (e.g., selectively attending to spelling-to-sound mappings) are necessarily subsumed by the achievement of proficiency (e.g., correctly and quickly pronouncing many words).

Reading research is relevant to public health

Research characterizing the behavior and neuroscience of reading is enriched by contributions from a panoply of methodologies, including lesion studies from Dejerine (1891, 1892) to Damasio and Damasio (1983); chronometric analyses from Cattell (1886) to Balota and colleagues (2004); and brain imaging experiments from Petersen and colleagues (1988) to Price and Devlin (2011). From a public health standpoint, the utility of reading research is

underscored by the fact that developmental dyslexia is common, accounting for 80% of cases of learning disorders (Lerner 1989), and affecting 5-10% of the population (Siegel 2006). Dyslexia is a specific disorder of reading acquisition in the face of adequate intelligence, instruction and motivation (some criteria for dyslexia include difficulties with spelling, e.g., http://www.ninds.nih.gov/disorders/dyslexia/dyslexia.htm). Notably, dyslexia poses significant barriers to educational success (Peterson and Pennington 2012) that persist into adulthood (Shaywitz and Shaywitz 2005). Because the neural basis of the disorder remains a matter of active research, with several competing or perhaps complementary hypotheses regarding its causes (see Ramus and Ahissar 2012 for a review), there is substantial room for improvement in the remediation of the disorder (Gabrieli 2009). Understanding reading in healthy subjects, especially with regard to its normal development, is almost certainly a prerequisite for establishing more effective approaches to treating dyslexia.

Chronometry is a window to neural processing

Chronometry, or the measurement and analysis of reaction times (RTs), is a pillar of psychological inquiry (Donders 1969; Sternberg 1969), and chronometric studies facilitated the earliest systematic investigations of the phenomenology of reading (Cattell 1886). However, the apparent simplicity of chronometry belies some of the more nuanced issues related to its proper application and interpretation. Measuring reaction times for overt vocal responses raises a particular methodological issue; specifically, the sensitivity of voice-activated keys in naming tasks varies with the physical shape of the onset phoneme of the pronounced item. Per one study (Spieler and Balota 1997), as much as 35% of the variance in response latencies during naming tasks is driven by this surface-level stimulus characteristic. If onset phonemes are not taken into account, for example by coding them (Treiman and Kessler 1995) or by orthogonalizing their distribution, the variance attributable to them may be assigned erroneously to experimental variables, inflating the presumptive effects of interest. (The pronunciations of

multi-syllabic words are much less uniquely influenced by onset phoneme characteristics (Yap et al. 2009)). In recognition of this potential confound, the stimulus sets used in this thesis work were chosen such that onset phonemes did not systematically vary by condition. Furthermore, the reaction times for the fMRI-associated data were computed by hand, since the waveforms in that study were unavoidably contaminated by scanner noise, and we wanted to detect response onsets with as much fidelity as possible. Taking pains to protect the accuracy and precision of reaction time measurements is particularly important for chronometric studies of lexical processing, because some of the effects of interest are quite small, on the order of tens of milliseconds, for example.

In addition to concerns about reaction time measurement, there are issues related to the analysis of reaction times as indices of behaviors of interest. In particular, reaction time analyses that consider only condition- and task- specific means (rather than multi-parameter descriptions of the reaction time distributions) are subject to omitting or even obscuring potential observations (e.g., Ratcliff 1979; Balota et al. 2008). Despite this caveat, and consistent with much of the psycholinguistic literature, the analyses in this thesis relied upon mean reaction times, calculated after purging for the very outlier trials that may, in fact, carry interesting information. The decision to use only the means to sample the underlying reaction time distributions for this thesis work arose partially by default, as a way of reining in the dimensionality of an already complex dataset. However, and as will be discussed in Chapter 4, subsidiary analyses involving ex-Gaussian distributions, for example, could be conducted in the future, in order to test more directly some of the conclusions suggested by the data presented in these chapters.

One final note about response latencies: this work focuses on reaction time rather than accuracy in the behavioral data for several reasons (note that accuracy data is nonetheless reported). First, there were potential ceiling affects for accuracy, especially for the adults

performing the read task; indeed, performance was generally high across both tasks and both age groups. Furthermore, errors could theoretically derive from any of several sources, sources that are impossible to distinguish *post hoc,* and therefore not accessible to a straightforward interpretation. For example, when a participant failed to regularize an irregular word (task described later), it was not possible to distinguish whether the person had (1) lapsed out of task mode; (2) failed to recognize that the item had an irregular spelling-to-sound correspondence; or (3) tried but failed to access the appropriate sublexical mappings in order to generate the response. It is for these reasons (and for the sake of brevity, in the case of Chapter 2) that when I talk about "behavior" in these pages, I am generally referring to reaction time. However, and as will be elaborated upon in Chapter 4, it should not be forgotten that accuracy is a very important component of task performance, with evidence existing to suggest that even small differences in error rate can produce relatively large differences in reaction time (Pachella 1974).

Functional neuroanatomical models of reading emphasize phonology, orthography and semantics

The notion that separate neural processing routes underlie different aspects of reading was originally advanced by Dejerine in the late 1800s as a result of his studies of patients presenting with focal brain lesions and dissociable reading deficits (Dejerine 1891, 1892). Even today, lesion studies continue to render useful insights about reading (Damasio and Damasio 1983; Henderson 1986; Fiez et al. 2006; Turkeltaub et al. 2013), providing particular utility with regard to constraining computational models (Nickels et al. 2008). Technological advances over the last several decades have given rise to increasingly sophisticated neuroimaging techniques, including magneto-encephalography (MEG) (Papanicolaou et al. 1998), event-related potentials (ERP) (Kutas and Hillyard 1980), positron emission tomography (PET) (Petersen et al. 1988; Petersen et al. 1990) and functional magnetic resonance imaging (fMRI) (McCarthy et al. 1993; Hinke et al. 1993). Imaging studies have certainly advanced our understanding of the neural

basis of reading, and in 2013 it has become difficult to keep up with the large number of such studies that are published every year. As Petersen and Fiez cautioned early in the development of PET studies, neuroimaging studies are especially useful when carefully designed, analyzed and interpreted in the context of a consideration of converging evidence from chronometric and lesion studies (1993). Indeed, one of my distinct chores in attempting to become a scholar in this field has been to think critically about how the observations that are rendered by one methodology affirm, contradict or simply coexist with observations deriving from another methodology, as the mappings between two "reading" studies (say, a modeling paper from the Coltheart group and a paper examining reading-related brain regions using functional connectivity) are not always readily accessible.

Based on early functional imaging data and certainly reflecting the behavioral literature as well, Pugh and colleagues have suggested a cognitive neuroscience model of reading development in which the functional contributions of left dorsal posterior regions, including parts of the left angular gyrus and left supramarginal gyrus, predominate as a child begins to read, allowing for analytic (slow, attention-requiring) integration of phonological, orthographic and semantic features of words (Pugh et al. 2001). As skilled reading develops, activity in left dorsal posterior regions trains a set of left ventral posterior regions, including the putative visual word form area (Cohen et al. 2002), allowing for fluent (fast, automatic) recognition of whole words. In this model, left anterior language regions, including Broca's area in the left inferior frontal gyrus, are responsible for phonological recoding and articulation (Greenlee et al. 2004), sharing some phonology-related response properties with regions in dorsal posterior regions (Church et al. 2011). The Pugh et al. model thus implies the primacy of an effortful, serial type of decoding – presumably relying heavily on phonological processing – for early reading, and emphasizes the importance of non-phonological recognition for more skilled reading. The model also summarizes the essential conceptualization of the neuroanatomy of reading that is embraced by

the field today, as reflected in both recent (e.g., Price 2012; Taylor et al. 2012) and slightly older (e.g., Jobard et al. 2003; Mechelli et al. 2003) meta-analyses.

Computational models of reading emphasize phonology and orthography

Neuroanatomical accounts of reading such as that of Pugh and colleagues have been guided and constrained in part by computational models that attempt to describe the mapping between visual orthographic input and speech output. Although this thesis work does not directly adjudicate between computational models, models of reading have so greatly influenced the way that people have thought about and studied reading that it is worth briefly describing their major features. Two classes of computational models in particular have historically garnered the most attention and empirical support in the reading literature. One class proposes two routes, while the other is rooted in a connectionist architecture.

In dual-route models, a sublexical (also referred to as phonological) route applies abstract spelling-to-sound rules (e.g., *b* is pronounced as /b/, *ea* is pronounced as /i/) while the second, more direct route (called lexical, or whole-word), associates rapidly recognized spellings with stored lexical forms (e.g. *cat* is recognized as /kæt/) (e.g., Coltheart et al. 1993; Coltheart et al. 2001). Stored lexical representations are strengthened in proportion to the exposure to a word's written word form, possibly instantiated as an increase in a node's baseline activation (Morton 1969) or a decrease in the threshold required for activation (Andrews 1989). Following phonological and lexical pathway processing, one response must be selected from among the 2 pathways' computations, with spelling-sound regularity determining whether the outputs of the 2 routes converge. It is notable that there is frequently no direct treatment of the final "phoneme buffer" or "response selection" stage (Pritchard et al. 2012) in dual route models, leading me to consider this final step in the reading process largely "neglected."

In contrast to dual route models, connectionist models posit distributed associations between orthographic, phonological, and semantic processes that emerge as the result of repeated exposure (training) to a finite corpus of words (e.g., Seidenberg and McClelland 1989; Plaut et al. 1996; Kello and Plaut 2003). Rather than being rule-based, the strengths of the associations between the various units change over time in response to co-activation. One criticism of connectionist models is that they seem little concerned with neuroanatomy, given their distributed nature. One attempt to map a connectionist model onto the brain was proposed by McCandliss and Noble (2003), and it posits a cascading mechanism whereby successful phonological decoding in the left posterior superior temporal gyrus helps to establish a rapid word recognition system in posterior regions (i.e., the visual word form area). Decoding skill, in this model, requires simultaneous attention to letters and to the sounds those letters make.

There exists an automaticity dogma in the reading literature

While the overwhelming contribution to the field of computational models of reading has been positive, I believe that there may be ways in which the simplifying assumptions of both connectionist and dual route models have unwittingly restricted the conceptual creativity of the empirical investigations that have accompanied the models' development. Specifically, the models have tended to focus on describing the component processes of reading (e.g. phonological, orthographic and semantic analysis), assuming that particular stimulus and task manipulations can be mapped onto these subcomponents, generally without considering the role of attentional control in shaping these processes and their integration (see McCann et al. 2000; Reynolds and Besner 2006 for further discussion). The empirical literature has largely reflected this focus on investigations of orthographic, semantic and phonological processing, with explicit investigations of the role of attentional control in reading many fewer in number and influence. My evaluation of the focus of the literature is that there exists something of an automaticity dogma throughout the field. To be clear, my opinion on this particular shortcoming

of the models is by no means an overall indictment, as I fully recognize their substantial and

often catalytic role in moving certain areas of investigation forward.

Automaticity is a thorny concept, requiring the consideration of many definitional criteria

(e.g., Stanovich 1990; Logan 1997). Key features of automatic processes include the speed and

effortlessness with which they are carried out, as well as the obligatory and/or subconscious

nature of their completion (Logan 1997). Descriptors such as fast and obligatory indeed

describe aspects of reading in skilled adults, and several seminal cognitive neuroscience

studies provide support for the notion that reading, in adults, shows features of automaticity.

From the Stroop effect (Stroop 1935), to Neely's demonstration of sematic priming (Neely 1977),

to Reicher's word superiority effect (Reicher 1969), these foundational studies are beautiful and

compelling. Compelling though they may be, studies that demonstrate automaticity with regard

to one component of reading processing under one specific set of experimental conditions do

not constitute general proof that the process of reading is outside of the purview of top-down

attentional control, even when a constellation of such studies is considered. Of particular

relevance to this work, whether or not skilled reading in adults can properly be considered

automatic, reading in a child is certainly neither fast nor effortless, and the trajectory by which

reading become facile remains in many respects a mystery.

There exists a phonology dogma in the reading literature

In case it is not adequately provocative to assert that the literature reflects a relative

paucity of interest in the role of attentional control in reading, I further propose that the field

additionally reflects a rather substantial bias towards thinking and talking almost obsessively

about phonology. A "strong phonological theory of visual word recognition" was laid out explicitly

by Ram Frost in 1998 (1998), but he was neither the first nor the last investigator to take the

position that reading is substantially concerned with phonology[3]. The phonology dogma may be especially apparent in the developmental literature, in which dyslexia research factors heavily, as there is solid evidence that children who experience reading difficulty demonstrate particular trouble in parsing the sounds of speech (Liberman et al. 1989; Ramus et al. 2003; Shaywitz and Shaywitz 2005).

Phonological processing, even in the unimpaired reader, is undoubtedly a crucial inroad to extracting the meaning from a printed word, particularly for readers of alphabetic orthographies[4] such as English. One could furthermore make an unassailably pragmatic argument that studying phonology is broadly useful to the reading investigator, since phonology is one of the shared features of written and spoken language. However, many people in the world read logographic scripts (consider the Chinese), in which symbols map directly onto meanings, rather than sounds. Not surprisingly, evidence suggests that Chinese reading skill hinges more on orthographic than phonological awareness (Tan et al. 2005), a fact that underscores the need to avoid equating reading with phonological processing. As was argued above with regard to the automaticity dogma, danger lurks in fidelity to a singular conceptual stance, and my own experience with this thesis work was that shifting my thinking from phonology to attention proved to be pivotal, a point that I hope will become self-evident.

The choice of reading task is of critical importance

The empirical data that has been amassed in tandem with the development of dual route and other computational models has derived primarily from lexical decision and speeded naming tasks. In lexical decision tasks, subjects decide whether a stimulus is a word or not; in

---

[3] Consult the October 2012 issue of *Behavioral and Brain Sciences* for an updated version of Frost's theory of reading (Cohen et al. 1990), as well as several peer commentaries on the topic, including a call by Kathleen Rastle (Frost 2012) to rethink the emphasis on phonology in reading theories.

[4] In alphabetic orthographies, the symbols of the writing system map onto sounds, with a variable degree of transparency between the grapheme-to-phoneme correspondences.

speeded naming tasks, subjects read words aloud as quickly and accurately as possible. Many

important observations have derived from other tasks as well. Indeed, one of the sources of

richness (and complexity) in the visual word recognition literature is the seemingly endless

variety of experimental paradigms that can be exploited to target particular aspects of word

reading. However, as Alecia Vogel showed as part of her thesis work in our laboratory, different

tasks are different (Vogel et al. 2012), and it is critical to think and speak precisely about the

particular processing requirements of a given task.

Read task

Overt speeded naming has both limitations and advantages as task used to study

reading. In contrast to lexical decision and similar judgment tasks (e.g. semantic classification),

naming does not involve an explicit decision process, and is thereby arguably relatively simpler

to decompose. However, speeded naming does require overt pronunciations, therefore

necessitating the preparation of an articulatory program followed by its execution (Cattell 1886).

Not only must one be careful to consider the motor aspects of word pronunciation when

conceptualizing a task involving an overt vocal response, but one could also make the argument

that aloud pronunciation of isolated words lacks somewhat in ecological validity, at least for

skilled adults, for whom the bulk of the reading in our daily lives involves the silent apprehension

of connected text. While the ecological validity complaint is not without merit, it is also the case

that in all areas of cognitive neuroscience, some level of reductionism is pragmatically inevitable

if one wishes to study a phenomenon in the laboratory.

Without appealing to pragmatics, I find it relatively easy, as the parent of an almost 5-

year-old, to justify the use of speeded naming as "ecologically valid" with regard to early skill

development. When a child is learning to read, sounding words out is indeed a very physical act,

and the engagement in the motor and auditory experience of overtly assembling the phonemes

as graphemes are encountered is something that persists throughout the early years of reading

instruction. Even adults, under circumstances requiring increased concentration, may revert to the physical act of forming the sounds of the letters with their mouths when reading, presumably reflecting the tight linkage between the motor and cognitive components of language. That being said, I certainly hope that none of you is currently reading this document out loud because you find it so utterly difficult to follow.

Turning from the behavioral literature on reading to the comparatively younger neuroimaging literature, the set of "reading-related" tasks that people have used seems even more motley. Because of worries about imaging signal artifact related to the motion of verbal articulation and the putatively puzzling physics of the oropharyngeal cavity, many PET and fMRI studies of reading have employed implicit tasks that require either a button-press response or not response at all. For example, silent reading may be coupled to an orthographic or phonological feature detection task, and participants are trusted to perform the assigned task faithfully (Turkeltaub et al. 2003). Other popular paradigms include rhyme judgment, syllable counting, word spelling, delayed verbal recall, and pseudohomophone discrimination (e.g., Rumsey et al. 1997; Joubert et al. 2004; Booth et al. 2004; Fiez et al. 2006). Each of these surrogate tasks requires computations above and beyond word identification, a fact that should be recognized when one considers the behavioral and imaging effects that they purport to reveal.

Regularize task

The regularize task has already been mentioned without having been described. In brief, the task is borrowed from a study conducted by Dave Balota, Jason Zevin and Mark Law (2000), and it involves the presentation of single stimuli (regular words, irregular words and pseudowords) and the instruction to "sound out" each item as though reading it for the first time. The regularize task thus challenges participants to direct attention exclusively to phonological processing, and to suppress their "sight-word" knowledge. Relegating to a later discussion the

specifics of the Balota et al. findings, suffice it to say that the task generated some very intriguing behavioral results, and despite this fact, very few subsequent studies have utilized the regularize task (but see Gold et al. 2005 for an imaging study using the regularize task). The reasons to adopt the regularize task in this work as a variant of speeded naming that requires top-down attentional control were manifold.

First, we sought to replicate the interesting behavioral results in young adults, an objective that is both important for keeping the field honest and that also seemed to be relatively low-hanging fruit. Moreover, the regularize task grabbed my attention as well-suited to a developmental study because of its apparent juxtaposition of something with which children should have great facility (i.e., sounding words out) vs. something with which children might struggle (i.e., inhibiting a prepotent response). Furthermore, the regularize task was appealing because its inputs and outputs (other than the irregular words) should be identical to those used in the read task, allowing for the possibility of closely targeting the processes that produce that mapping from stimulus to response. Lastly, aside from potential concerns about an MRI speaking artifact, the regularize task seemed to be begging to be taken into the scanner, ripe for an exploration of the neural activity underlying its performance (note that the Gold et al. study cited above (2005) used a very different analysis approach from that which we use here). One legitimate criticism of the regularize task is that it may be different enough from normal reading so as to be relatively uninformative for generating insights into how the attentional control of normal reading occurs. I think this concern is not unfounded, and I certainly want to be cautious about overreaching my conclusions. Nonetheless, it seems unlikely to me that the fundamental attentional mechanisms that are brought to bear on the regularize task are somehow wildly irrelevant to the control of reading under "normal" circumstances.

<u>**Reading research: issues pertinent to development**</u>

  The second part of this chapter addresses several themes that are especially relevant to the study of reading development.

Dyslexia research highlights the need to pay more attention to attention

  The importance of attentional mechanisms in the development of normal reading ability is suggested by at least one line of research in the developmental dyslexia literature. Although a clear etiological and mechanistic description of dyslexia remains elusive, there is compelling evidence that deficits in visuo-spatial attention contribute to difficulty in learning to read, at least in some children (e.g., Franceschini et al. 2012; see Vidyasagar and Pammer 2010; Valdois et al. 2004 for reviews). Visual attention span deficits have been observed in dyslexic children across languages (e.g., Bosse et al. 2007), with one interesting interpretation of such deficits positing that dyslexics are "sluggish" in their ability to shift attention (Hari and Renvall 2001), possibly accounting for the range of sensory-related deficits (Hari et al. 2001; Laasonen et al. 2000; Habib 2000; Hari et al. 1999; Helenius et al. 1999; Eden and Zeffiro 1998; Stein and Walsh 1997) that have been observed as part of the dyslexic phenotype. Relatedly, dyslexia co-occurs with attentional deficit hyperactivity disorder (ADHD), particularly the inattentive subtype, much more frequently than what would be expected by chance (e.g., Willcutt and Pennington 2000; August and Garfinkel 1990), suggestive of common underlying genetic factors (Paloyelis et al. 2010) and possibly a mutual cognitive deficit (Castellanos and Tannock 2002; McGrath et al. 2011). Although this thesis work does not address dyslexia directly (and indeed, potential participants were screened for reading difficulties), it is certainly possible to imagine ways in which the studies presented here could be adapted to examine dyslexia.

Development is orthogonal to skill acquisition

In these pages, "age-group differences" connotes those improvements in reading skill that accompany increased experience with reading, improvements which correlate with age for our study population. This shorthand is not meant to imply that getting older *per se* confers increased reading skill. Indeed, as will be alluded to in Chapter 3, one of the challenges inherent to interpreting developmental data is the disambiguation of "age group" differences due to developmental processes (e.g., prefrontal cortical maturation, conferring improvements in cognitive control) from experience-dependent processes specific to the acquisition of a skill (e.g., the consolidation of lexical representations of words, facilitating more rapid access of particular items from the lexicon). It should also be mentioned that the 8-10 year old children who participated in the study described in Chapter 3 were reading at a significantly higher grade level than suggested by their chronological ages. Generalization from the present sample of participants to the entire population of "typical" 8-10 year old children must therefore be made cautiously, if at all.

Reaction times generally decrease across development

Consistent with what most of us would surmise based on observation, experimental studies confirm rather unequivocally that children are slower and less accurate than young adults (Elliott 1970) on just about any cognitive task (Kail 1991). Processing speed, measured as reaction time, decreases robustly from childhood through adolescence and into adulthood, with some evidence for domain specificity in the trajectories of reaction time improvement across development (e.g., Kail and Miller 2006; but see Hale 1990). Although it is likely that developmental differences in such "noncentral" factors as incentive, motivation, attentiveness and practice contribute to the reaction time differences observed in children vs. adults (Wickens 1974), most investigators agree that the bulk of the observed age group differences arise from fundamental limitations of information processing and capacity in the developing brain.

Identifying those fundamental limitations of processing using behavioral assessments has proven somewhat difficult, however. One attempt at decomposition that is particularly interesting, in light of the present results, referred to Smith's stages of stimulus processing (Smith 1968) and concluded that response selection (rather than stimulus input, stimulus categorization or response execution) most distinguishes improvements in reaction time across one particular range of development, from age 5 to 11 years (Fairweather and Hutt 1978).

Brain maturation involves significant development of attentional control

Behavioral and neuroimaging evidence converges to suggest that brain maturation involves important and identifiable changes in cognitive/ executive/ attentional control processes (e.g., Zelazo et al. 2004) presumably contributing greatly to the performance discrepancy between children and adults discussed above. Although a full treatment of this topic is outside of the scope of this thesis, the essential theory is that the comparatively slow development of particular brain regions, including parts of prefrontal cortex (Casey 1997; Casey et al. 2000), has the consequence of making children relatively less able than adults to exert control over their processing systems. Evidence suggests that there are substantial regional-specific changes (e.g., Huttenlocher and Dabholkar 1997) in control-related brain regions into adolescence (Gogtay et al. 2004) and even beyond (Luna et al. 2001). Children in the age range of 8-10 years old, like those who participated in the study in Chapter 3, are worse than young adults in terms of their abilities to switch from one task to another (Crone et al. 2006); monitor their own performance (Kopp 1982); alert appropriately to stimuli (Rueda et al. 2004); inhibit prepotent responses (Davidson et al. 2006); and utilize working memory (Luna et al. 2004).

Road map to the data chapters

Both data chapters in this these use a comparison, borrowed from Balota, Law and Zevin (2000), between a simple speeded reading task vs. a regularize ("sound out") task. The first data chapter focuses on young adults only, and examines both behavioral and imaging data. The second data chapter focuses on school-age children and young adults, leveraging behavioral data only.

In Chapter 2, I discuss the results of a combined behavioral/neuroimaging study in 2 young adult cohorts. I first show that we were able to replicate the essential findings of the Balota et al. study, specifically confirming that lexicality and frequency influence read and regularize task performance differently, underscoring the susceptibility of reading to modulation by top-down control. Consistent with the interpretation of Balota and colleagues, I furthermore argue that the reaction time patterns observed in the 2 young adult cohorts are consistent with 2 distinct mechanisms by which top-down attentional control interacts with reading processes, pathway control and response checking. I then present evidence, motivated by the 2-mechanism hypothesis, that 2 sets of brain regions, together including members of Dorsal attention (Corbetta and Shulman 2002), Cingulo-opercular and Frontal-parietal (Dosenbach et al. 2006) control networks show separable activity patterns that map nicely onto roles reflecting pathway control and response checking.

In the Chapter 3, I show that 8-10 year old children, like young adults, can perform the regularize task. Although children made more mistakes on the regularize task than did the adults, the early readers were faster than the experienced readers to regularize, and this speed advantage for children held for both words and pseudowords. Because children are slower than adults across a range of cognitive tasks (e.g., Kail 1991) – with children showing particular immaturity with regard to inhibiting prepotent responses (e.g., Davidson et al. 2006) – the developmental observation is remarkable in and of itself. The results of several analyses

considering interactions of several stimulus characteristics (lexicality, frequency and regularity) with task and age group are then presented, in an attempt to identify the causes of the robust crossover interaction. I argue that those analyses both *generally* reconfirm the plausibility of the 2-mechanism model of the attentional control of reading, and *specifically* suggest that both pathway control and response checking are incompletely mature in less experienced readers. *Post hoc* supplemental analyses, including subject- and item- level correlations of task and neuropsychological measures, bolster the twin conclusions of Chapter 3.

Together, Chapters 2 and 3 suggest that dissociable subcomponents of attentional control interact with subcomponents of reading processing, and that these interactions are dynamic across skill development and across task demands.

<u>Summary</u>

Part of the objective of this thesis work is to echo the suggestion of Reynolds and Besner (2006) that the field of reading would do well to question the possibly obfuscating conceit of automaticity, making room for novel insights that are not readily accessible within theoretical constructs that ignore the modulatory potential of attentional control. At the very least, it seems important to unpack just what is meant by automaticity and attentional control when it comes to reading. My take on the literature at this point is that there is plenty of room for more pointed investigations of *where* – in terms of both boxes/sticks and voxels/neurons – attentional control systems interact with reading processes.

Let us, then, dive into searching for the locus or loci of the attentional control of reading, both in terms of cognitive models and the brain. As one of my favorite investigators in this field once opined,

"[I]t is incumbent upon the researcher

to determine the locus or loci

of the effects being observed."

- Balota and Chumbley (1985)

CHAPTER 2: SEPARABLE ROLES FOR ATTENTIONAL CONTROL SUB-SYSTEMS IN READING TASKS: A COMBINED BEHAVIORAL AND fMRI STUDY

S. Katie Ihnen, Steven E. Petersen and Bradley L. Schlaggar

Abstract

Attentional control is important both for learning to read and for performing difficult reading tasks. A previous study invoked *two* mechanisms to explain reaction time (RT) differences between reading tasks with variable attentional demands (Balota et al. 2000). The present study combined behavioral and neuroimaging measures to test the hypotheses that there are *two* mechanisms of interaction between attentional control and reading; that these mechanisms are dissociable both behaviorally and neuroanatomically; and that the two mechanisms involve functionally separable control systems. First, RT evidence was found in support of the two-mechanism model, corroborating the previous study (Balota et al. 2000). Next, two sets of brain regions were identified as showing fMRI BOLD activity that maps onto the two-mechanism distinction. One set included bilateral Cingulo-opercular regions and mostly right-lateralized Dorsal Attention regions (CO/DA+). This CO/DA+ region set showed response properties consistent with a role in reporting which processing pathway (phonological or lexical) was biased for a particular trial. A second set was composed primarily of left-lateralized Frontal-parietal (LFP) regions. Its signal properties were consistent with a role in response checking. These results suggest that subcomponents of attentional control interact with subcomponents of reading processes in healthy young adults.

Introduction

The cultural development of reading is one of humankind's most revelatory accomplishments. Despite its importance in modern life, fluent reading is something that many people take for granted, perhaps because it, like many other learned skills, is achieved with a remarkable degree of automaticity across many contexts (e.g., Neely 1977; Schneider and Shiffrin 1977). Fluid and automatic reading is certainly not taken for granted by the 5-10% (or more) of the population that is estimated to suffer from dyslexia, or impaired reading despite normal intelligence and adequate instruction (Siegel 2006). The link between attention and reading ability is underscored by the observation that dyslexia co-occurs with attention deficit hyperactivity disorder (ADHD) more frequently than would be expected by chance, with comorbidity estimates ranging from 18-45% (August and Garfinkel 1990; Willcutt and Pennington 2000; Willcutt et al. 2010; Germano et al. 2010). Additionally, evidence suggests that deficits in visual attention span may contribute to dyslexia independently of phonological processing skills (for reviews, see Valdois et al. 2004; Vidyasagar and Pammer 2010).

Automatic processes (like skilled reading in typical adult subjects) tend to be fast; relatively effortless; proceeding to some degree obligatorily; and incompletely accessible to conscious awareness (Logan 1997; but see also Stanovich 1990). Descriptors such as fast and obligatory indeed describe aspects of reading, evidenced prominently by an entire literature exploring variants of the classic Stroop effect (Stroop 1935; Tanenhaus et al. 1980; Dennis and Newstead 1981; see MacLeod 1991 for a review). Non-Stroop studies of adult readers have furnished additional evidence for automatic access to various word codes during reading tasks, as shown, for example, by the intrusion of orthographic information during tasks emphasizing phonology (e.g., Perfetti 1992; Booth et al. 1999; Booth et al. 2008) or by the observation that word meanings can be extracted from brief stimulus presentations, even if the meanings cannot be reported explicitly (Luck et al. 1996). Evidence that reading task variants are prone to

interference from irrelevant processing features underlies one of the critical questions in the

reading literature and, indeed, a primary motivation for the current study: If one of the features

of skilled reading is automaticity, can experienced readers nonetheless exert selective

attentional control over the component processes of reading? If so, how is such attentional

control achieved?

One study suggested that *two distinct* mechanisms must be invoked to explain fully the

reaction time differences between a simple reading task and an effortful (attention-requiring)

reading task (Balota et al. 2000). Balota, Law and Zevin compared a speeded naming (read)

task vs. a novel regularize task, using a dual route framework. In contrast to the read task, the

regularize task required subjects to pronounce words based exclusively on spelling-to-sound

principles, pronouncing "pint" (/pajnt/) to sound like "lint" (/lInt/). The authors reasoned that if

skilled readers can direct attentional control *completely* to *frequency-independent* grapheme-to-

phoneme conversion to regularize words, then the typically observed main effects of lexicality

and word frequency (suggestive of lexical route activation) should be reduced in the regularize

task as compared to the read task.

Several key findings were reported by Balota and colleagues (2000). First, and not

surprisingly, subjects performed the regularize task more slowly than the read task. Secondly,

participants showed greater slowing to regularize words than pseudowords, losing completely

the positive lexicality effect (speed advantage for words) they showed in the read task. The

reliable interaction of task by lexicality suggested that skilled readers had trouble completely

inhibiting lexical (whole-word) processing for words, despite task instructions. Lastly, the usual

word frequency effect (high frequency words are pronounced more quickly than low frequency

words) *reversed* in the regularize task for regular, but not irregular words, producing a significant

interaction of task, frequency and regularity. This final observation was interpreted as evidence

of the need for more than one mechanism of attentional control. As the authors wrote, "the

reversal of the word frequency effect [because it is selective for regular words] cannot simply be attributed to increased interference from the lexical processing pathway…" (2000).

Specifically, Balota and colleagues suggested that lexical route suppression – a concept closely related to the pathway control hypothesis invoked to explain other experimental findings (Baluch and Derek 1991; Monsell et al. 1992; Rastle and Coltheart 1999; Zevin and Balota 2000) – and a subsequent "verification procedure" are *both* important mechanisms of interaction between reading processing and attention during the regularize task. The verification procedure, or response checking, can accommodate the word frequency effect reversal for regular words if, perhaps, on some portion of the high frequency regular word trials, an additional time penalty is incurred because the participant second-guesses the (normal) pronunciation, subconsciously thinking the pronunciation should sound "funny." A related post-lexical, pre-articulation step has been described by Balota and Chumbley in the context of a lexical decision task as the "neglected decision stage" (1984).

The discussion in the Balota et al. study (2000) of a 2-mechanism interaction between reading and attentional control – pathway control in the form of lexical route suppression, plus response checking – was framed within a dual route model. According to dual route reading theory, visually presented words are processed along two distinct routes, the phonological and the lexical (e.g., Coltheart et al. 1993; Coltheart et al. 2001). The phonological route concatenates grapheme-to-phoneme mappings in a serial manner (Forster and Davis 1991) that is independent of word frequency and word regularity, i.e., "chunking" sound together, regardless of task and stimulus features (Church et al. 2008). Pseudowords can only be completely decoded via the phonological route. The lexical route, on the other hand, involves accessing whole-word representations of stimuli that have been learned through repeated exposure (e.g., Visser and Besner 2001), thus resolving more quickly for words that are more frequently encountered (Frederiksen and Kroll 1976), and without regard to word regularity

(Coltheart et al. 2001). Word regularity determines whether the outputs of the 2 routes converge (regular words) or not (irregular words), a distinction that cannot be made until the "full reading process" occurs, i.e., after route computation (Taylor et al. 2012). Critically, "phonology is always partly assembled and always partly lexical," even if each of those processing route outputs remains incompletely specified (Coltheart et al. 2001; see also Frost 1998), as, for example, with pseudowords. Prior to response articulation, "phoneme recognition" occurs (Pritchard et al. 2012), i.e., a response is selected from between the two possibilities generated by the routes.

Dual route models are computationally capable of accommodating the types of shifts in response time and accuracy due to top-down attentional control that Balota and colleagues noted in their regularize paradigm (Rastle and Coltheart 1999), even if such "strategy" effects have most certainly not been a focal point of investigation with regard to the models (see Reynolds and Besner 2006 for a more detailed discussion). Empirical studies have manipulated stimulus context (and used other kinds of paradigms (Paap and Noel 1991; Zevin and Balota 2000)) in order to encourage subjects to increase their reliance on one type of processing over another. For example, when people read lists composed mainly of pseudowords, they make more regularization errors and show delayed RTs for exception (~irregular) words (Baluch and Derek 1991). Context-dependent performance differences, including those arising from list composition manipulations, have often been interpreted as evidence for what is referred to as the pathway control hypothesis, referenced earlier (e.g., Pugh and Rexer 1994; Monsell et al. 1992; Rastle and Coltheart 1999; Zevin and Balota 2000; Reynolds and Besner 2005; but see Lupker et al. 1997; Jared 1997 for conceptual frameworks other than pathway control). Attentional effects have also been noted at the response checking/selection stage. For example, compared to naming tasks, lexical decision tasks show greater effects of whole-word variables like word frequency (Balota and Chumbley 1985; Balota and Chumbley 1990; Balota et al. 2004;

but see Monsell et al. 1989 for a contrasting interpretation). In sum, ample evidence suggests

not only that there are indeed situations in which reading requires attentional control, but also

that there are likely multiple loci (Vogel et al. 2005) at which attention may influence reading

(including pathway control and response checking).

Just as reading involves several component processes that might serve as points of

interaction with attention control, attentional control itself is achieved through the dynamic

interaction of several brain systems (Dosenbach et al. 2008; Petersen and Posner 2012).

Dissociation of control systems is demonstrable through converging methods, including the

observation of measurable discrepancies, within an individual, in scores on tests thought to

reflect the function of different aspects of attentional control (Fan et al. 2002). Separability of

control systems is additionally supported by both task based fMRI (e.g., Corbetta and Shulman

2002; Dosenbach et al. 2006; Spreng et al. 2010), functional connectivity MRI (e.g., Seeley et al.

2007; Dosenbach et al. 2007; Vincent et al. 2008; Power et al. 2011; Yeo et al. 2011) and lesion

(e.g., Nomura et al. 2010) data at the group level, showing distinct patterns of activity and

consistent network relationships between sets of regions, for example Cingulo-opercular,

Frontal-parietal and Dorsal attention networks.

Figure 2.1 depicts a unified model, combining aspects of the dual route theory and

principles expressed by Petersen and Posner (2012), of how attentional control may interact

with reading processing in skilled adult readers. The model proposes 2 points of interaction. An

early point of interaction ("Mechanism #1") involves setting the bias to emphasize processing

along either the lexical or the phonological route in a task- and stimulus- dependent manner

("pathway control"). The second point of interaction ("Mechanism #2") entails the selection and

verification of the response from between the 2 route outputs ("response checking"). The model

accounts for the primary behavioral effects reported by Balota et al. (Balota et al. 2000) and

replicated in the present study (a task-dependent lexicality effect and an impact of word

regularity on the task reversal of the word frequency effect), while also constraining novel

predictions regarding how control-related brain regions may interact with reading-related brain

regions.

Figure 2.1 Schematic depiction of two-mechanism model of attentional control interacting with reading processes
       Following presentation of a visual word stimulus (i.e., plaid) and basic perceptual processing, activation automatically proceeds along 2 processing routes. The phonological route is shown on top (/p l e d/) and the lexical route is depicted on the bottom (*plaid*). Following route processing, a single response is selected and prepared for articulation (? /pled/ or /plæd/ ?), a step that some theorists call "phoneme recognition" (see (Pritchard et al. 2012) Figure 1, for an updated dual route schematic). Finally, an overt response is articulated. Mechanism #1 (pathway control) marks the proposed site of early interaction between an attentional control system(s) and reading processing, reflecting the biasing of processing along one of two routes, depending on task demands. Mechanism #2 (response checking) marks the proposed site of a later interaction between an attentional control system(s) and reading processing, to ensure the appropriate response is selected from among potentially competing alternatives.
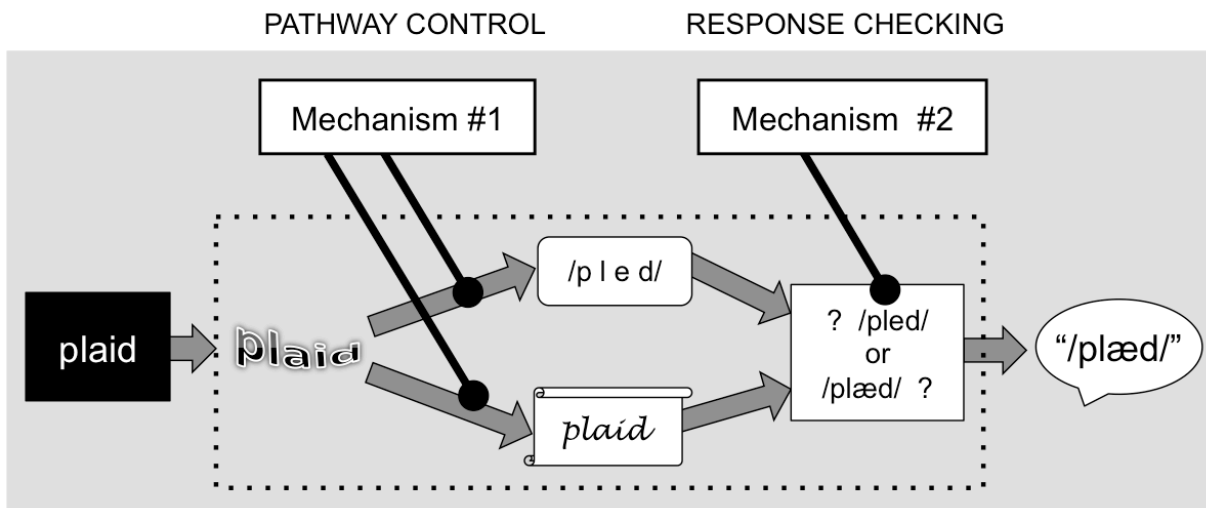
Specifically, the 2-mechanism model predicts that there will be at least 2 different kinds

of control signals, anatomically segregated from one another and from data-processing systems

(Posner and Petersen 1990). Control signals consistent with a role at each mechanism in the

model should map plausibly onto previously defined functional roles. For example, Cingulo-

opercular (CO) control regions, thought to be important for parameter setting and maintenance of task mode (especially with regard to their task-sustained responses) (Dosenbach et al. 2007), may play a role at Mechanism #1, reflecting or reporting the pathway control required for a particular task/stimulus combination. Regions of the Frontal-parietal (FP) control network may act at Mechanism #2, subserving response checking, consistent with moment-to-moment "adaptive control" (Dosenbach et al. 2007; Dosenbach et al. 2008). It is possible that brain regions belonging to other attentional control networks might additionally/alternatively be involved in interacting with reading processing regions in the proposed model. For example, resting state functional connectivity fMRI (rs-fcMRI) work from our laboratory has demonstrated a privileged functional relationship between regions in the Dorsal attention system and reading-related regions including the putative Visual Word Form Area (VWFA) (Vogel et al. 2012; see also Zhao et al. 2011).

Despite the fundamental connection between attention and reading, the regularize paradigm used by Balota et al. has remained little studied over the last decade, particularly in terms of neuroimaging studies (but see Gold et al. 2005). The present study employed a comparison of a read task vs. a regularize task in 2 independent samples, a Behavioral cohort ("Study 1") and an fMRI cohort ("Study 2"). We tested the hypotheses that there are 2 mechanisms of interaction between attentional control and reading processes (pathway control and response checking); that these mechanisms are dissociable both behaviorally and neuroanatomically; and that the two mechanisms involve functionally separable control systems.

Materials and Methods: General

The methods common to both the behavioral and neuroimaging studies are described first, followed by the methods unique to each study.

Participants

Participants were recruited from the Washington University community, and were screened for the following: neurologic and/or psychiatric diagnoses; history of dyslexia or reading difficulty; and current use of psychotropic medications. Imaging study subjects were additionally screened for MRI contraindications (e.g., metal implants, pregnancy, cardiac pacemakers).

Enrolled participants for Study 1 (Behavioral cohort) included 32 right-handed, native and monolingual English speakers (ages 22-28 years; 15 male). Of the 32 enrolled Study 1 participants, all met the task accuracy criterion (>65%) for inclusion. Thus data from all 32 were included in the analyses. All recruitment, screening, and experimental protocols for both studies were approved by the Washington University Human Studies Committee. Subjects provided written informed consent and were compensated for their time.

35 participants (19 male) ages 21-28 years were enrolled in Study 2 (fMRI cohort). Two participants did not complete the imaging session due to personal reasons, and one participant performed the task too slowly to render separable hemodynamic responses for each stimulus. Of the final set of 32 Study 2 participants, one subject did not contribute RT data, due to equipment failure. However, the accuracy of that subject's responses was monitored online, so that the BOLD data are usable and therefore included. The final data set includes data from 32 subjects (31 with RT data).

Neuropsychological tests were administered to each participant. Two subtests of the Wechsler Abbreviated Scale of Intelligence (Vocabulary and Matrix Reasoning) assessed general intelligence (Wechsler 1999). Three subtests of the Woodcock–Johnson III (Letter–Word ID, reading Fluency, and Word Attack) estimated absolute and relative reading level (Woodcock and Johnson 2002). The Elision subtest of the Comprehensive Test of Phonological

Processing (CTOPP) assessed phonological processing ability (Wagner et al. 1999). A Stroop color-naming task assessed attentional control (Spieler et al. 1996).

Table 2.1 shows the group mean scores (with standard deviations) for neuropsychological tests for subjects from both studies. All subjects were reading at or above the 25th percentile, and all had IQs at or above the mean. As shown, the average scores on all 3 subtests of the Woodcock Johnson reading assessments were high, reflecting a sample of highly skilled readers. Similarly, Verbal IQ estimates placed both study cohorts into the above average range.

Table 2.1 Neuropsychological testing results

Neuropsychological test or subtest	Study 1 M (SD)	Study 2 M (SD)
Wechsler Abbreviated Scale of Intelligence (WASI) IQ, percentile	97.26 (4.45)	92.53 (11.61)
WASI Vocabulary subtest[a], ss	16.66 (1.47)	15.50 (2.13)
WASI Matrix Reasoning subtest, ss	13.94 (1.05)	13.25 (1.68)
WJ Letter/Word Identification[b], ss	118.44 (7.76)	115.25 (10.10)
WJ Reading Fluency, ss	121.13 (16.47)	114.81 (17.95)
WJ Word Attack, ss	102.41 (9.25)	102.06 (8.45)
Comprehensive Test of Phonological Processing (CTOPP), Elision subtest, raw	19.06 (1.08)	18.16 (2.95)
Stroop facilitation effect[c] (raw RT)	2 ms (54)	18 ms (26)
Stoop interference effect[c] (raw RT)	86 ms (51)	94 ms (52)

Note. [a]Population range for each WASI subtest is 1-19. [b]Population mean for each WJ subtest is 100. [c]For Study 2 cohort, Stroop data was lost for 1 subject, so averages include only 31 subjects.
ss = scaled score

Stimuli

Study 1 included 636 stimuli, comprising 252 pseudowords (pronounceable, nonsense letter strings) and 384 words. The words included 96 of each of 4 combinations of regularity (regular or irregular) x frequency (high or low) and length (short or long). Short items were 3-5 letters (mainly 1 syllable) and long items were 6-8 letters (mainly 2 syllables). Primary analyses for Study 1 were conducted using short items only (n = 318), consistent with both Balota et al. (2000) and with Study 2. See Table 2.2 for a summary of the important lexical properties of Study 1 short stimuli, and Appendix B for more detailed stimulus information, including a complete list of stimuli (Supplemental material B.1). Long (2-syllable) items were not the focus of the current analyses, nor was a direct examination of length effects. However, the interested reader is referred to Supplemental material B.4 to see the mean values of lexical properties for the 5 long stimulus types, and Supplemental table D.6 to see the reaction time and accuracy rates for the Study 1 cohort, considering long items only. Generally, the statistical analyses of interest yielded similar results for the 1- and 2- syllable items (data not shown).

Study 2 included 320 stimuli, 272 of which overlapped with the short items from Study 1. For Study 2, all items were short (3-5 letters; mainly 1 syllable). The stimulus set included 128 pseudowords and 192 words. The words include 48 of each of 4 combinations of regularity (regular or irregular) x frequency (high or low). See Table 2.3 for a summary of the important lexical properties of Study 2 stimuli, and Appendix B for more detailed stimulus information, including a complete list of items in Supplemental material B.1. Note that the imaging analyses of Study 2 collapsed across frequency.

Table 2.2 Mean values of lexical properties for 5 short stimulus type: Study 1

Stimulus type[a]	Number of letters (SD)	Number of syllables (SD)	Number of orthographic neighbors[b] (SD)	Number of phonological neighbors[b] (SD)	Standard frequency index[c] (SD)
High frequency regular words	4.25 (0.64)	1.00 (0.00)	8.15 (4.89)	15.81 (9.35)	59.51 (5.12)
Low frequency regular words	4.25 (0.64)	1.00 (0.00)	7.46 (5.09)	15.52 (9.91)	46.49 (3.87)
High frequency irregular words	4.33 (0.56)	1.00 (0.00)	6.23 (4.01)	15.52 (9.09)	61.50 (5.96)
Low frequency irregular words	4.29 (0.65)	1.04 (0.20)	5.65 (5.03)	15.27 (10.91)	46.55 (3.60)
Pseudowords	4.29 (0.62)	1.00 (0.00)	5.34 (4.48)	---	---

Note. [a]For all word stimulus types, n = 48; for pseudowords, n = 126.

Table 2.3 Mean values of lexical properties for 5 short stimulus type: Study 2

Stimulus type[a]	Number of letters (SD)	Number of syllables (SD)	Number of orthographic neighbors[b] (SD)	Number of phonological neighbors[b] (SD)	Standard frequency index[c] (SD)
High frequency regular words	4.40 (0.57)	1.00 (0.00)	7.69 (4.38)	15.08 (8.61)	61.91 (5.47)
Low frequency regular words	4.40 (0.61)	1.00 (0.00)	7.17 (5.05)	15.06 (10.34)	46.58 (3.93)
High frequency irregular words	4.40 (0.57)	1.02 (0.14)	6.08 (4.45)	15.71 (9.46)	62.51 (6.35)
Low frequency irregular words	4.33 (0.63)	1.10 (0.31)	5.04 (4.73)	13.85 (10.50)	46.59 (3.31)
Pseudowords	4.37 (0.59)	1.00 (0.00)	5.20 (4.43)	---	---

Note. [a]For all word stimulus types, n = 48; for pseudowords, n = 128.

Note for Tables 2.2 and 2.3: [b]Orthographic neighbors and phonological neighbors calculated using the English Lexicon Project database (http://elexicon.wustl.edu; Balota et al. 2007). [c]Standard frequency index (SFI) is from Zeno et al. (Zeno et al. 1995).

Tables 2.2 and 2.3 show that stimuli were matched, as much as possible, across the conditions of interest for lexical properties that were not explicitly manipulated. Although not indicated in the tables, care was also taken to ensure that acoustic properties of the onset phoneme did not systematically vary by stimulus type (Spieler and Balota 1997). For both studies, stimuli were randomly intermixed within task runs, and each subject either performed the read task or the regularize task on each stimulus.

Word frequency ratings were based on the Zeno et al. standard frequency index (SFI) (Zeno et al. 1995); stimuli spanned an SFI continuum from 40.0 to 75.0 (unit = log-transformation of a weighted frequency per million) (see Supplemental figure B.8 for a depiction of the distribution of items across the frequency range). A median split designated subsets (regular and irregular) of words as high frequency (M SFI = 60.1) and low frequency (M SFI = 47.0). High-frequency words had significantly higher SFIs than low-frequency words in each study (Study 1: t(165*) = 20.34; p < .001; Study 2: t(157*) = 22.12; p < .001; see Tables 2.2 and 2.3). Additionally, frequency was *matched* across regularity classes for both studies. That is, regular and irregular words *within* each frequency category were not statistically different from one another in terms of SFI: Study 1 low: t(94) = 0.07; p = 0.95; Study 1 high: t(94) = 1.76; p = 0.08; Study 2 low: t(94) = 0.01; p = 0.99; Study 1 high: t(94) = 0.50; p = 0.62.

A word was deemed regular (rather than irregular) if its rime (e.g., –at in cat) followed typical spelling-to-sound correspondences (n.b., there is no fixed set of spelling-to-sound rules in English). For example, "gave" (/gev/) is regular, whereas "have" (/hæv/) is irregular, because the "e" at the end of "have" should generate the pronunciation /hev/. Pilot testing confirmed that those words designated as irregular typically resulted in different pronunciations for the 2 tasks, while regular words did not. Pseudowords were created *de novo*, using onset (e.g., /k/ in cat) and rime (e.g., /æt/ in cat) graphemes that tend to be both consistent and regular (i.e., unambiguously pronounced).

The stimuli were divided into fixed lists: 6 lists of 106 items each for Study 1, and 10 lists of 32 items each for Study 2. Stimulus types were equally distributed across the lists. For both studies, the list ordering and list task assignment were counterbalanced across participants.

Task instructions, response collection and accuracy scoring

For the regularize task, subjects were instructed to "sound out" items (real words and nonsense words) as though reading them for the first time, using whatever strategies they normally use when encountering new items. Subjects were given 13 demonstration trials as well as 73 practice trials (13 demonstration stimuli + 60 new stimuli) for the regularize task (see Supplemental material B5 for practice items). The practice session was self-paced with feedback; correct answers were provided when needed. In Study 2, the practice session occurred prior to the start of scanning, out of the scanner, and a slightly different list of practice stimuli was used (see Supplemental material B6 for practice items). None of the practice items from either study appeared as stimuli in the experimental runs for that study.

Overt vocal responses were coded by a single rater for accuracy of pronunciation. For the regularize task, a pronunciation of a regular word (e.g., "hike") was correct if it corresponded to the standard dictionary pronunciation. An irregular word was correct if its pronunciation obeyed any of several potentially phonologically legitimate alternative spelling-to-sound mappings of its rime (e.g., pronouncing "have" as /hev/). For both tasks, pseudowords were considered correct if their pronunciations followed standard orthographic-phonological mapping conventions, or rhymed with similarly spelled real words (e.g., /krot/ for "crote").

Responses were recorded so that reaction times could be extracted offline. Raw sound files (one per subject per run) were recorded for each run for each subject, using a Sony MP3 Digital Voice Recorder (ICD-UX70 1GB USB) in Study 1, and Cool Edit 2000 software for Study 2 (Syntrillium). Recorded files were normalized, noise-reduced and purged for artifacts (e.g.,

coughs, throat clearings) using Audacity, an open-source software for sound editing (http://audacity.sourceforge.net/).

For Study 1, reaction times were calculated from cleaned files using an in-house Matlab program (Matlab 7.80; R2009a; The MathWorks, Natick, MA) (Nelles et al. 2003). RTs were also spot-checked manually using Audacity. In contrast to Study 1, trial-by-trial reaction times for Study 2 were calculated *manually* by a single rater, using Audacity. Manual RT determination was necessary because some of the subjects' responses coincided with the scanner noise (i.e., did not occur during the 1 s gap in the sequence), rendering the automated reaction time extraction program unreliable.

To account for possible individual differences in general processing speed, trial-by-trial RTs for both studies were *z*-score transformed (Faust et al. 1999). For brevity, the reaction time results presented for both studies are discussed only in terms of *z*RT, but essentially the same pattern of results was observed when untransformed reaction times were used instead.

The mean and standard deviation were calculated on an individual subject basis across all correct trials of both tasks, excluding raw reaction times < 200 ms or > 3 standard deviations above a subject's mean as probable outliers (on average, fewer than 1.5% of trials per subject). Only correct trials were included in the reaction time and BOLD analyses. Repeated-measures analyses of variance (rmANOVAs) were planned at the subject-level to examine effects of task and lexical variables (including lexicality, word regularity and word frequency); additional analyses at the item-level in Study 1 confirmed that the pattern of results was not driven by the particular selection of stimuli (data not shown). Appropriate post hoc analyses clarified the results of the planned comparisons.

Following Balota et al. (2000), the interaction analyses of the reaction time data including frequency and lexicality were computed using only regular words and pseudowords. Since both

regular words and pseudowords generate the same pronunciations for the 2 tasks, their task comparison is free of potentially confounding articulatory effects.

Materials and Methods: Study 1 (Behavioral)

Task structure

For Study 1, 3 consecutive runs of the regularize task were followed by 3 runs of the read task ("Pronounce these items normally"). Each run included 106 items and lasted 6 min, 10 s. Pilot testing suggested that the fixed regularize/read ordering was most effective for establishing and maintaining task mode.

Stimuli were displayed on a black background in lowercase white letters at font size 56, subtending 2.1 to 5.0 degrees of visual angle horizontally and 0.7 to 1.1 degrees vertically. Each stimulus appeared alone, and in the middle of the screen, for 2500 ms (see Supplemental figure A.1). Participants were free to respond as soon as the stimulus appeared and up until the next stimulus appeared. Immediately following the disappearance of each stimulus a white fixation crosshair appeared in the middle of the screen for 1 s. Participants were instructed to maintain visual fixation on the screen for the duration of the run.

Apparatus

Stimuli were controlled using Psyscope X B53 (Cohen et al. 1993) on a Power Macintosh G3, and were displayed on a 16 in Sony CRT computer monitor. A standing microphone was placed 2-3 cm away from the participant's mouth, and transmitted spoken responses to the Sony MP3 Digital Voice Recorder.

Materials and Methods: Study 2 (Behavioral + fMRI)

Task structure

Following standard localizer and anatomical scan acquisition (described below), each

participant performed 5 runs each of the two tasks, in the following sequence: 2 runs of the

regularize task, 2 runs of the read task, 2 runs of the regularize task, 2 runs of the read task,

and finally one run each of the regularize and the read task. This quasi-alternating run order

contrasts with the blocking order utilized in Study 1, and was selected to balance the conflicting

objectives of minimizing the need for task-switching vs. safeguarding against losing an entire

task (e.g., as a block of runs) due to subject movement, equipment failure, etc. Each run

included 32 items and lasted 5 min, 21 s. Trials were jittered with an intertrial interval of 1, 2 or 3

frames. The scan time required for the task runs averaged 75 min.

During the scan session, most subjects underwent 2 additional 5 min 30 s runs of resting

state fMRI (rs-fcMRI) data collection. One of these runs typically occurred after the first 4 task

runs, and the second run typically occurred after the last task run. Those data are not discussed

further.

Stimulus presentation

To discourage movement during scanning, subjects were fitted with an individualized,

thermoplastic mask that was attached to the head coil, serving chiefly as a landmark for head

position rather than a restraint, *per se*. A mirror was also placed on top of the head coil, allowing

visualization of an LCD projection screen near the caudal end of the tube. Stimuli were

projected onto the screen from the adjacent control room, using Psyscope X B53 (Cohen et al.

1993) on a Macintosh iMac. Stimuli were displayed on a black background in lowercase white

letters at font size 56, so that the stimuli subtended 2.1 to 4.9 degrees of visual angle

horizontally and 0.7 to 1.05 degrees vertically. Each stimulus appeared alone, and in the middle

of the screen, beginning at 1350 ms after the onset of an image acquisition frame (see

Supplemental figures A.1 and A.2). The stimulus remained on the screen for 1550 ms. This

timing was chosen, based on pilot studies, in order to maximize the likelihood that vocal

responses would occur within the 1000 ms gap in scanning that occurred after each TR, as

discussed below. During the 1-3 frames (jittered) between trials, a white fixation crosshair

appeared in the middle of the screen. Participants were free to respond as soon as the stimulus

appeared and up until the next stimulus appeared (i.e., within 6-12 s). Participants were

instructed to maintain visual fixation on the crosshair for the duration of the run. Subjects were

fitted with an MRI-compatible headset and microphone, with the microphone positioned 1-2 cm

away from the mouth.

Image collection

Images were acquired using a Siemens 3T Trio scanner (Erlangen, Germany) and a

Siemens 12-channel Matrix head coil. A single high-resolution structural scan was acquired

using a sagittal magnetization-prepared rapid gradient echo (MP-RAGE) sequence (slice time

echo = 3.08 ms, TR = 2.4 s, inversion time = 1 s, flip angle = 8 degrees, 176 slices, 1 x 1 x 1

mm voxels). Functional runs were acquired parallel to the anterior-posterior commissure plane

using an asymmetric spin-echo echo-planar pulse sequence (TR = 2.0 s plus a 1.0 s delay, total

TR = 3.0 s; T2* evolution time 27 ms; flip angle 90 degrees). 32 contiguous interleaved 4mm

axial slices, with 4 x 4 mm in-plane resolution, allowed for total brain coverage. Magnetization

steady state was assumed after 12 s, so that functional data acquisition began with the fifth MR

frame. As mentioned, the pulse sequence used here included a 1 s delay after each TR to

facilitate the collection and monitoring of subjects' overt responses.

Image preprocessing

Despite the requirement for overt vocal responses, in-scanner movement was relatively

low. Frame-by- frame movement correction data from the rotation and translation in the x, y, and

z planes was computed for each subject for each run, and no runs had overall movement greater than 1.00 mm RMS. Average per-run movement was 0.24 mm RMS.

Automated image preprocessing included the following procedures: (1) removal of a single pixel spike caused by signal offset; (2) correction of odd vs. even slice intensity differences due to interleaved slice acquisition; (3) debanding; (4) quantification of and correction for movement within- and across- runs, via realignment of slices into scanner space using rigid-body rotation and translation; (5) correction for magnetic field distortions using subject-acquired field maps; and (6) within-run normalization of signal intensity to a whole-brain mode of 1000.

Preprocessed functional BOLD data from each subject were next registered to a common atlas using 12-parameter affine warping of the individual's MP-RAGE to the target. The target atlas was based on Talairach and Tournoux (Talairach and Tournoux 1988), and was created in-house[5] by mutually co-registering the anatomy of 12 healthy young adults and 12 healthy school-age children (Brown et al. 2005). As part of the atlas transform, the data were resampled on an isotropic 2 mm grid.

Next was performed a set of preprocessing steps that have been shown to be useful (Power et al. *in* preparation) for minimizing motion-related noise in the BOLD signal in resting state MRI analyses (e.g., Fox et al. 2009; Power et al. 2012; Satterthwaite et al. 2013), versions of which have been implemented in task based fMRI data (Zhao et al. 2011). The complete sequence included demeaning and detrending each run followed by across-run regression of the following: (1) six parameters obtained by rigid body head motion correction; (2) the whole brain signal averaged across the entire brain; (3) ventricular signal averaged from ventricular

---

[5] For the purposes of reporting, all coordinates from these data have been converted to MNI space. Coordinates from other studies are reported in the original space used by the authors, along with an MNI conversion (http://www.sdmproject.com/utilities/?show=Coordinates).

ROIs; and (4) white matter signal averaged from white matter ROIs. The first derivatives of these regressors were also used, yielding a total of 18 regressors.

Supplemental material A.3c describes further the motivation for applying this final preprocessing step and additional details about the methodology, and also explains that the fundamental results were similar when analyses were recomputed using data that did not undergo this step (see also Supplemental figure A.3).

Scope of imaging analysis

Primary effects of task and lexicality, secondary effects of regularity

The hypothesis of a two-mechanism interaction between attentional control sub-systems and reading processing implies testable predictions regarding task, lexicality, and regularity (description of predictions to follow; also see Table 2.4). Region selection was thus driven by identifying, in a voxelwise manner, regions showing (1) an interaction of task by timecourse, and/or (2) an interaction of task by lexicality by time. Effects of regularity were then explored in secondary analyses using the task- and task x lexicality- sensitive regions.

Effects of task and lexicality were prioritized for region definition (rather than effects of regularity) for two reasons: they showed the largest and most reliable behavioral effects, and they could be computed using the full stimulus set (i.e., words and pseudowords), therefore being more likely to detect even small effects. Figure 2.2 depicts schematically the critical pieces of the analysis stream for Study 2, the steps of which are next described in detail.

Figure 2.2 Analysis flow for Study 2, imaging study

First (A), whole-brain ANOVAs identified individual regions showing task- and task x lexicality effects. Next (B), the timecourses of the regions showing the 2 effects of interest were computed. Then (C), concatenated timecourses were used to cluster the regions using a UPGMA algorithm. Clusters were mapped (D) to the cortical surface for visualization, along with putative community borders. For each cluster (E), effects of task, lexicality and regularity were interrogated, by computing rmANOVAs treating regions as "subjects" and clusters as the grouping variable. The results of the rmANOVAs were used to test the predictions outlined in Table 2.4.
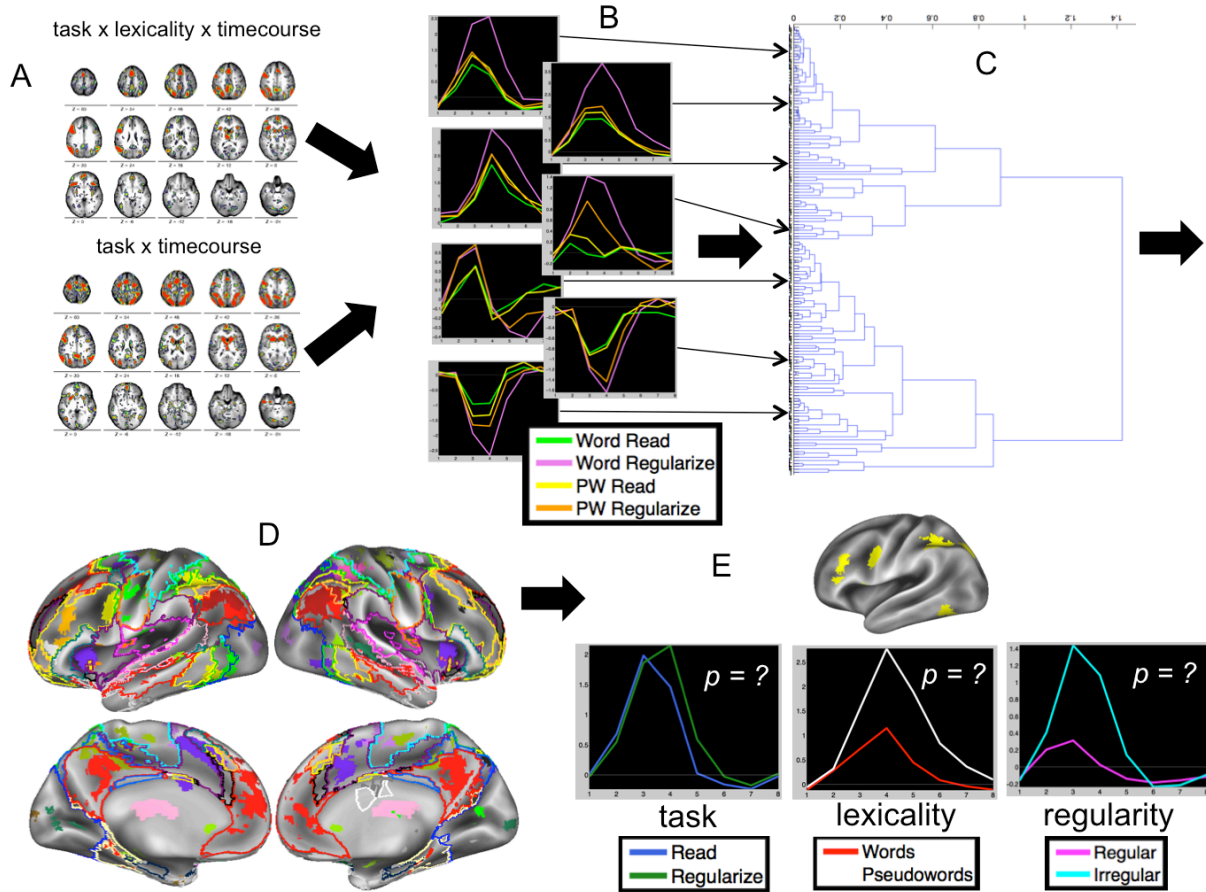
Image analysis: region identification

GLM modeling

Statistical analyses of the preprocessed BOLD data were computed at the voxel level using a general linear model (GLM) (Ollinger et al. 2001). The GLM design incorporated time as an 8-level factor, with the 8 levels corresponding to successive MR frames following presentation of the stimulus. Two levels of task (read or regularize) and several levels of stimulus type (including lexicality, as well as regularity and frequency for words) were also included as factors in the design matrix. Events of interest included only correct trials; errors were modeled separately in the GLM and were ignored here. No assumptions were made about the shape of the hemodynamic response function (HRF). Timecourses for all analyses were entered into ANOVAs using random effects models. GLM modeling and a portion of the subsequent analyses were conducted using in-house software programmed in the Interactive Data Language (IDL) (Research Systems, Inc. Boulder, CO). Additional statistical analyses were conducted using SPSS 16.0.1 (SPSS Inc., 2007) and Matlab (Matlab 7.80; R2009a; The MathWorks, Natick, MA).

Voxelwise image generation

Task x timecourse

First, a 2 task (read vs. regularize) by 8 timepoints (timecourse) voxel-wise whole brain rmANOVA was conducted. This image identified regions whose BOLD timecourses showed a deflection from baseline (positive or negative), that depended on which task the subject was performing.

Lexicality x task x timecourse

Next, a 2 lexicality (all words vs. pseudowords) by 2 task (read vs. regularize) by 8

timepoints (timecourse) voxel-wise whole brain rmANOVA was conducted. This image identified regions showing differential BOLD timecourse activity for the two tasks that *also* depended on whether the stimulus in a particular trial was a word or a pseudoword.

Thresholding and peak identification

The voxel-wise analyses described above produced images containing voxels showing a task x timecourse interaction or a lexicality x task x timecourse interaction. These two images were corrected for false positives, based on Monte Carlo simulation, using a criterion of 24 contiguous voxels with a $z > 3.5$ (Forman et al. 1995; McAvoy et al. 2001). Regions were then extracted from each of the two images using an in-house peak-finding algorithm courtesy of Avi Snyder. Activity peaks were identified within the Monte Carlo-corrected images by applying a smoothing kernel of 4 mm kernel and requiring that adjacent peaks be no closer than 10 mm apart.

Timecourse extraction

The task x timecourse and lexicality x task x timecourse rmANOVAs described above were recomputed, this time using the regions identified by the peak definition as ROIs, in order to derive the timecourses of activity for each subject, for every level of every factor in the design. Average timecourses of the BOLD % signal change at each of 8 timepoints were then computed and visualized for the various conditions.

Creation of primary region pool

Finally, regions from the 2 analyses ((task x timecourse) and (lexicality x task x timecourse)) were pooled into a single set of possible ROIs. Region sets were combined in order to cast as wide a net as possible in identifying regions that may contribute to the 2 most reliable features of the behavior (the task effect, and the modulation of the task effect by lexicality); also there was considerable overlap between the images. Overlapping regions and

regions with biologically implausible timecourses were excluded from the ROI pool as described in Supplemental material A.4.

Hierarchical clustering of identified regions

The 2-mechanism model depicted in Figure 2.1 predicts that 2 different types of attentional control signals interact with reading processing regions in the read/regularize task paradigm. The primary objective of the imaging analyses was therefore to test the hypothesis that subsets of attentional control-related regions from the overall pool (perhaps including members of systems such as the Frontal-parietal, Cingulo-opercular, Dorsal attention, etc.) would segregate into 2 distinct sets based on response properties, suggesting dissociable contributions to the tasks. All the regions in the final pool of task-and task x lexicality- sensitive regions were therefore segregated, based on BOLD responses, using hierarchical clustering.

Creation of dendrogram and visualization of region clusters

Hierarchical clustering was computed in Matlab (Matlab 7.80; R2009a; The MathWorks, Natick, MA). To focus on the rise and peak of the hemodynamic response, the first 6 timepoints (i.e., 18 s) of each average timecourse were isolated, for each of four conditions (2 lexicality x 2 task), for each region. (Thus timecourses for the regions originally identified in the task x timecourse image had to be re-extracted from a task x lexicality x timecourse rmANOVA.) The 4 timecourses for each region were concatenated, creating a matrix of 24 columns with as many rows as there were regions. The UPGMA (Unweighted Pair Group with Arithmetic Mean) method was used to generate clustering assignments between the regions represented in the matrix (e.g., Ploran et al. 2007; Church et al. 2011).

The dendrogram resulting from the UPGMA clustering was examined by plotting the observed clusters, at various similarity (1-r) thresholds, on the surface of the brain, using Caret software (Van Essen et al. 2001). The 2 clusters that appeared to fall into attentional control-

related network boundaries (i.e., putative attentional control clusters) were carried forward into subsequent analyses.

Predictions of imaging analysis: Mechanism #1 and Mechanism #2

Several predictions regarding the BOLD activity patterns at Mechanism #1 and Mechanism #2 guided the remaining analyses. The predictions, described below and outlined in Table 2.4, are designed to test (1) whether or not the region sets identified using hierarchical clustering show different functional properties and (2) whether or not the roles of the region sets map onto the model schematized in Figure 2.1.

Table 2.4 Predicted BOLD fMRI effects for regions acting at Mechanism #1 and Mechanism #2

Prediction		Mechanism #1 PATHWAY CONTROL		Mechanism #2 RESPONSE CHECKING	
Task effect		Small		Large	
Lexicality effect		"Opposite…………. and equal"		"Opposite……....... and unequal"	
	Read task	pseudowords > words	Small	pseudowords > words	Small
	Regularize task	words > pseudowords	Small	words > pseudowords	Large
Regularity effect	Both tasks	Small, irregular > regular		Large, irregular > regular	
General response properties		A relatively brief response, reflecting a bias signal		A relatively protracted response, reflecting a response check	
		Overall smaller BOLD responses		Overall larger BOLD responses	

Regions acting at Mechanism #1 may show signals reflecting top-down pathway control, for example a read-out as to whether a participant was directing attention to lexical or phonological processing in the context of a particular task and stimulus. If pathway control is

akin to parameter setting (in that it is required for every task regardless of difficulty (Meiran 1996)), regions acting at Mechanism #1 might be expected to show responses that are generally less variable across the 2 tasks in terms of peak % BOLD change. Such reporting signals may also be predicted to be relatively rapid to rise and rapid to fall. For example, the effects of lexicality (words vs. pseudowords) for the 2 tasks might be small, relatively transient, of roughly equivalent magnitudes and in opposite directions for the 2 tasks. The "opposite and equal" prediction for the lexicality effect reflects that each task requires a bias signal, but that those bias signals are different. I.e., skilled readers presented with familiar words presumably emphasize the lexical pathway during the read task, and the phonological pathway during the regularize task. Also, regions reflecting pathway control should show minimal effects of word regularity, consistent with the dual route postulate that both regular and irregular words are processed along both pathways essentially obligatorily (Coltheart et al. 2001).

Unlike pathway control, presumably necessary for any task variant, response checking is likely *more* critical for the task requiring overriding the usual stimulus-response mapping, regularize. If response checking matters more for regularize than for read, then Mechanism #2 regions should show generally larger effects of task than the early actors, both as a main effect and as interactions with other variables. Thus Mechanism #2 regions should show lexicality effects that are "opposite and unequal" (specifically, larger in the regularize task, where the response checking assumes a larger importance). Similarly, the effect of regularity at Mechanism #2 regions should be larger for the regularize task than the read task (in addition to being larger overall in Mechanism #2 regions compared to Mechanism #1 regions; see above). It is also likely that regions showing signals associated with response checking will show a more protracted timecourse of response than will regions that reflect pathway control, owing to the time required to complete a decision-making process.

To test this set of predictions, timecourses from each region in each UPGMA-generated cluster were extracted and entered into statistical analyses as described below.

Plotting average timecourses for each control-related cluster

Average lexicality x task timecourses plotted

Timecourses (extracted as described above) were averaged, across all regions of each of 2 putative attentional control clusters, for each of the 4 conditions originally leveraged to compute the clustering (lexicality (2) x task (2)). Cluster-average timecourses were then plotted to show which features of the BOLD responses across these 4 conditions may have driven the clustering. (Because these lexicality x task timecourses comprised the very data that entered into the clustering, clusters will, by definition, show different patterns across these timecourses.)

Regularity x task timecourses extracted, and average timecourses plotted

Timecourses were next extracted for all regions in the final set by computing a rmANOVA over 2 regularity (regular vs. irregular) by 2 task (read vs. regularize) by 8 timepoint (timecourse). As for task x lexicality, average timecourses across all regions of the 2 putative control clusters were then plotted for each of 4 combinations of regularity and task. Cluster-average timecourses were visualized in order to depict qualitatively whether clusters showed obvious distinctions in their sensitivity to regularity across the two tasks.

Quantification of differences between region clusters

To test quantitatively the imaging predictions, several rmANOVAs were next computed in which each *region* from each cluster was treated as "a subject" and each regions' *cluster assignment* was treated as a grouping variable. First, two omnibus rmANOVAs were computed, one that included as within-"subject" variables lexicality (2) x task (2) x timecourse (8), and a second that included regularity (2) x task (2) x timecourse (8). Next, a set of rmANOVAs was computed with cluster as the grouping variable and only timecourse (8) as the within-"subject"

variable, considering each of the 4 combinations of lexicality x task and each of the 4 combinations of regularity x task separately. Follow-up planned comparisons additionally assessed for separate effects of cluster membership on lexicality (2) x timecourse (8) for each task; task (2) x timecourse (8) for each lexicality; regularity (2) x timecourse (8) for each task; and task (2) x timecourse (8) for each regularity. This full complement of analyses (all of which featured cluster as the grouping variable) allowed for the identification of which of the many possible effects of lexicality, task and regularity best characterizes the functional segregation of the 2 putative attentional clusters, and in particular whether those differences map onto the hypothesized 2-mechanism model shown in Figure 2.1.

The rmANOVAs conducted using cluster as a grouping variable were not all equally unbiased. Specifically, the regions were originally defined as showing reliable effects of either task x timecourse or lexicality x task x timecourse, and the resulting clustering was computed using the concatenated timecourses from the 4 conditions of task (2) x lexicality (2). Therefore, the omnibus lexicality x task x timecourse x cluster ANOVA was *biased* towards a reliable result, although statistical significance was not a *necessary* consequence of the analysis. It is conceivable that timecourse differences causing 2 sets of regions to segregate in cluster space are not large enough and/or reliable enough to result in a statistically significant 4-way interaction. Importantly, the *post hoc* tests (e.g., examining the effect of task for words only) are unbiased, since they consider only portions of the original data used to define the clusters. Our interpretations emphasize the results of the *post hoc* tests.

In contrast to the lexicality analyses, the rmANOVAs using regularity as a factor are completely unbiased, as the regions were defined blind to regularity. The statistical reliability of the omnibus rmANOVA and all of the *post hoc* tests involving regularity can therefore be assessed without caveats related to bias.

Results

Key results from Balota et al.'s behavioral analysis were replicated (twice)

The reaction time data measured both out and in the scanner confirmed the generalizability of the key Balota et al. findings, driving a principled investigation of the associated neural activity in Study 2. In order to recapitulate most closely the Balota study and to keep the 2 studies of the current report comparable, primary reaction time analyses were conducted over the short items only. Reaction time analyses were conducted on zRT data (Faust et al. 1999) at the subject level; analyses on raw reaction times yielded similar results (data not shown). Table 2.5 shows the mean raw reaction time and accuracy rates for both Study cohorts for both tasks, for various stimulus types.

For brevity, statistical analyses of the accuracy data are not reported in this chapter; they largely mirrored the reaction time analyses. Also for brevity, statistical results are presented only for the main effects and interactions that were specifically targeted in order to replicate the Balota et al. findings and constrain and inform the imaging analyses. Chapter 3 provides a more complete treatment of the RT and accuracy results, in both adults and children.

Participants were slower to regularize than to read

Repeated measures ANOVAs were conducted separately for each study to examine the main effect of task, collapsed across all stimulus types. As shown in Figure 2.3A, in both studies, there was a significant main effect of task, driven by slower performance on the regularize than the read task (Study 1: mean RTs of 1170 ms vs. 733 ms; rmANOVA $F(1,31) = 348.45$, $p < .001$; Study 2: mean RTs of 1464 ms vs. 1036 ms; rmANOVA $F(1,30) = 463.58$, $p < .001$).

Compared to the Study 1 cohort, subjects in the scanner were slower to perform both tasks, by ~350 ms.

Table 2.5 Mean raw reaction times (ms) and accuracy rates, with SD (short items)

	Read				Regularize			
	Study 1		Study 2		Study 1		Study 2	
Stimulus type	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)
High frequency, regular	672 (119)	1.00 (.000)	965 (184)	0.999 (.007)	1242 (295)	0.989 (.025)	1556 (580)	0.988 (.021)
Low frequency, regular	691 (129)	0.999 (.007)	982 (182)	1.000 (.000)	1209 (295)	0.991 (.028)	1517 (596)	0.988 (.019)
High frequency, irregular	706 (125)	0.995 (.013)	994 (179)	0.984 (.028)	1252 (290)	0.856 (.122)	1554 (597)	0.912 (.080)
Low frequency, irregular	739 (129)	0.970 (.041)	1046 (230)	0.956 (.055)	1246 (251)	0.845 (.100)	1577 (596)	0.899 (.083)
Pseudo-words	783 (175)	0.990 (.016)	1094 (262)	0.992 (.016)	1081 (277)	0.990 (.022)	1342 (546)	0.993 (.013)

Direction of lexicality effect depended on task

Repeated measures ANOVAs were next conducted separately for each study using the factors of task and lexicality, including regular words and pseudowords only.

There was a significant interaction of task and lexicality in Study 1 ($F(1,31) = 109.12$, $p < .001$; see Figure 2.3B). This interaction was driven by a positive effect of lexicality in the read task (+101 ms; $F(1,31) = 37.61$, $p < .001$), and a negative lexicality effect in the regularize task (-145 ms; $F(1,31) = 60.26$, $p < .001$). That is, participants were *faster* to *read* regular words than pseudowords but *slower* to *regularize* words than pseudowords.

As in Study 1, there was a significant interaction of task and lexicality for subjects in the scanner: ($F(1,30) = 139.57$, $p < .001$; see Figure 2.3B. This interaction again reflected faster performance for words (than pseudowords) in the read task (+120 ms; $F(1,30) = 66.02$, $p < .001$), and slower performance for words in the regularize task (-195 ms; $F(1,30) = 52.61$, $p < .001$).

For both Studies, task and lexicality interacted significantly for the comparison of irregular words vs. pseudowords as well (both Fs > 70), showing a very similar pattern as that observed for regular words, i.e., positive lexicality effects for read and negative lexicality effects for regularize.

Figure 2.3 Behavioral results for the main effect of task and the interaction of task x lexicality

Raw reaction times, for both Study 1 and Study 2, for the main effect of task (A) and the interaction of task x lexicality for regular words (B). Raw RTs are shown; statistical analyses were conducted using the *z*-transformed RTs. For each pair of plots, the y-axis (in ms) has the same scale, though different values, reflecting the ~350 ms RT slowdown for participants in the scanner. Error bars show +/- 1 standard error of the mean (SEM). Because the statistical analyses were computed as repeated-measures ANOVAs, the error bars may be used only to assess group variability; they cannot be used to infer statistical significance. Asterisks indicate significant main effects and interactions, at the following *p* values: *** *p* ≤ .001; ** *p* ≤ .005; * *p* ≤ .05.
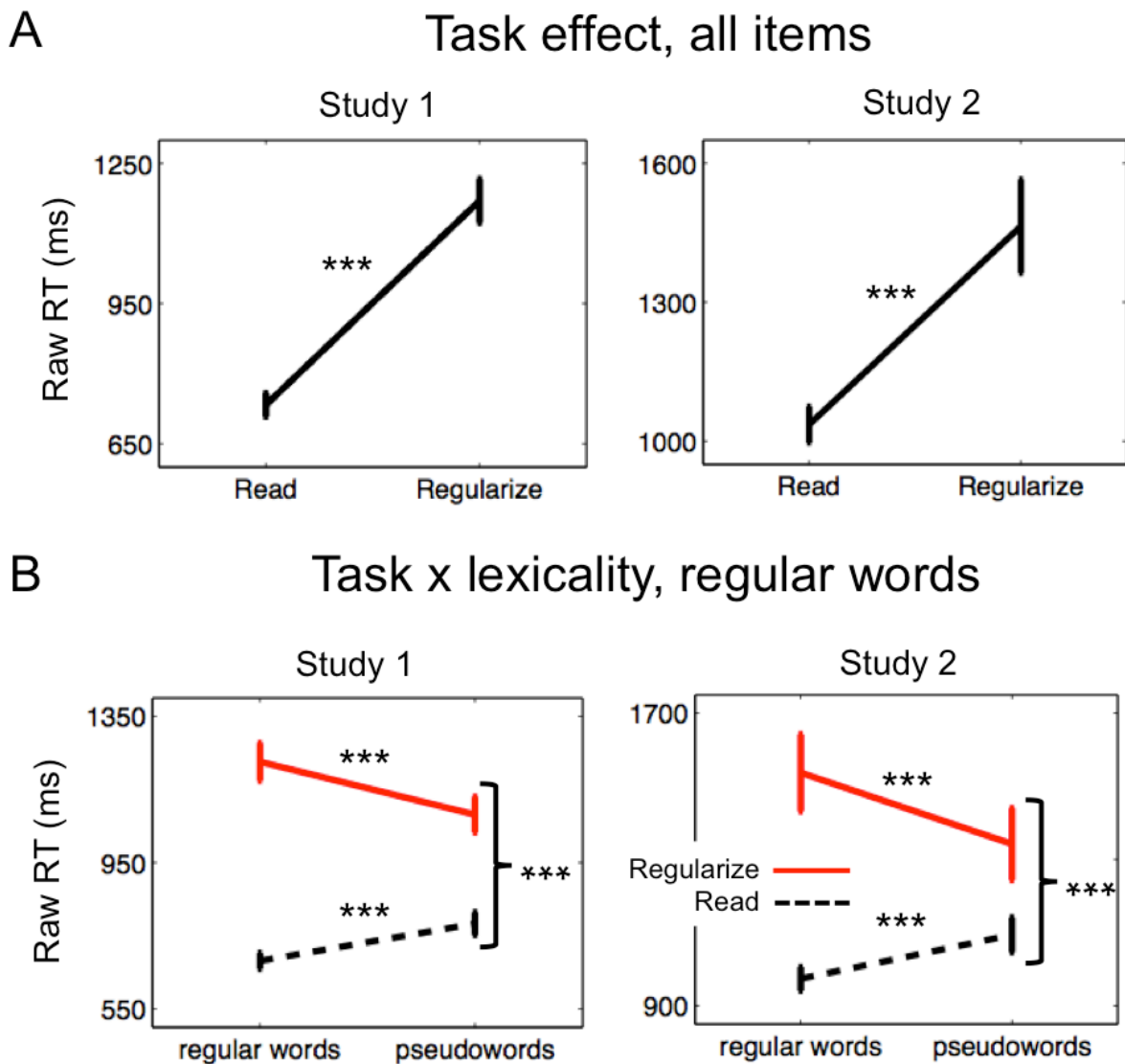
Word frequency effect reversal was obtained for regular words for regularize task

Repeated measures ANOVAs were next conducted separately for each study using the factors of task and frequency, including regular words and pseudowords only.

In Study 1, there was a significant interaction of task and frequency for regular words ($F(1,31) = 13.31$, $p = .001$), driven by a significant positive frequency effect of +19 ms in the read task ($F(1,31) = 9.51$, $p = .004$) and a significant negative frequency effect of -33 ms in the regularize task ($F(1,31) = 6.78$, $p = .014$). See Figure 2.4A.

As shown in Figure 2.4A, a significant interaction of task and frequency was obtained for the regular words in the Study 2 subjects as well ($F(1,30) = 9.01$, $p = .005$), driven by a positive (but non-significant) frequency effect of +17 ms in the read task ($F(1,30) = 1.15$, $p = .293$) and a significant negative frequency effect of -39 ms in the regularize task ($F(1,30) = 8.10$, $p = .008$).

Task x frequency effect for irregular words was inconsistent between Studies

Repeated measures ANOVAs were next conducted separately for each study using the factors of task, frequency and regularity. In contrast to Balota et al., in neither cohort was a significant interaction of task x regularity x frequency obtained (Study 1: $F(1,31) = 1.08$, $p = .307$; Study 2: $F(1,30) = 1.38$, $p = .250$).

To evaluate further the nature of the observed null 3-way interactions, rmANOVAs were next conducted separately for each study using the factors of task, frequency and regularity. Unlike what Balota et al. reported (Balota et al. 2000), the interaction of task and frequency reached significance for short irregular words in Study 1 ($F(1,31) = 5.11$, $p = .031$). As shown in Figure 2.4B, the interaction reflected a frequency effect for irregular words that was positive and significant for read (+33 ms; $F(1,31) = 20.45$, $p < .001$) and small, negative and nonsignificant for regularize (-6ms; $F(1,31) = 0.007$, $p = .933$).

Consistent with Balota et al., but inconsistent with Study 1, results from in the scanner confirmed a null interaction of task and frequency for irregular words ($F(1,30) = 1.06$, $p = .312$).

As shown in Figure 2.4B, the null interaction in Study 2 reflected a positive frequency effect for

irregular words for both tasks that was sizeable and significant for read and smaller and

nonsignificant for regularize (+51 ms for read; $F(1,30) = 16.79$, $p < .001$; +23 ms for regularize;

$F(1,30) = 1.96$, $p = .172$).

Figure 2.4 Behavioral results for the interaction of task x frequency, stratified by regularity
Raw reaction times, for both Study 1 and Study 2, for the interaction of task x frequency, separately for regular (A) and irregular (B) words. As in Figure 2.3, error bars show +/- 1 standard error of the mean (SEM), and asterisks indicate significant main effects and interactions, at the following p values: *** $p \le .001$; ** $p \le .005$; * $p \le .05$.
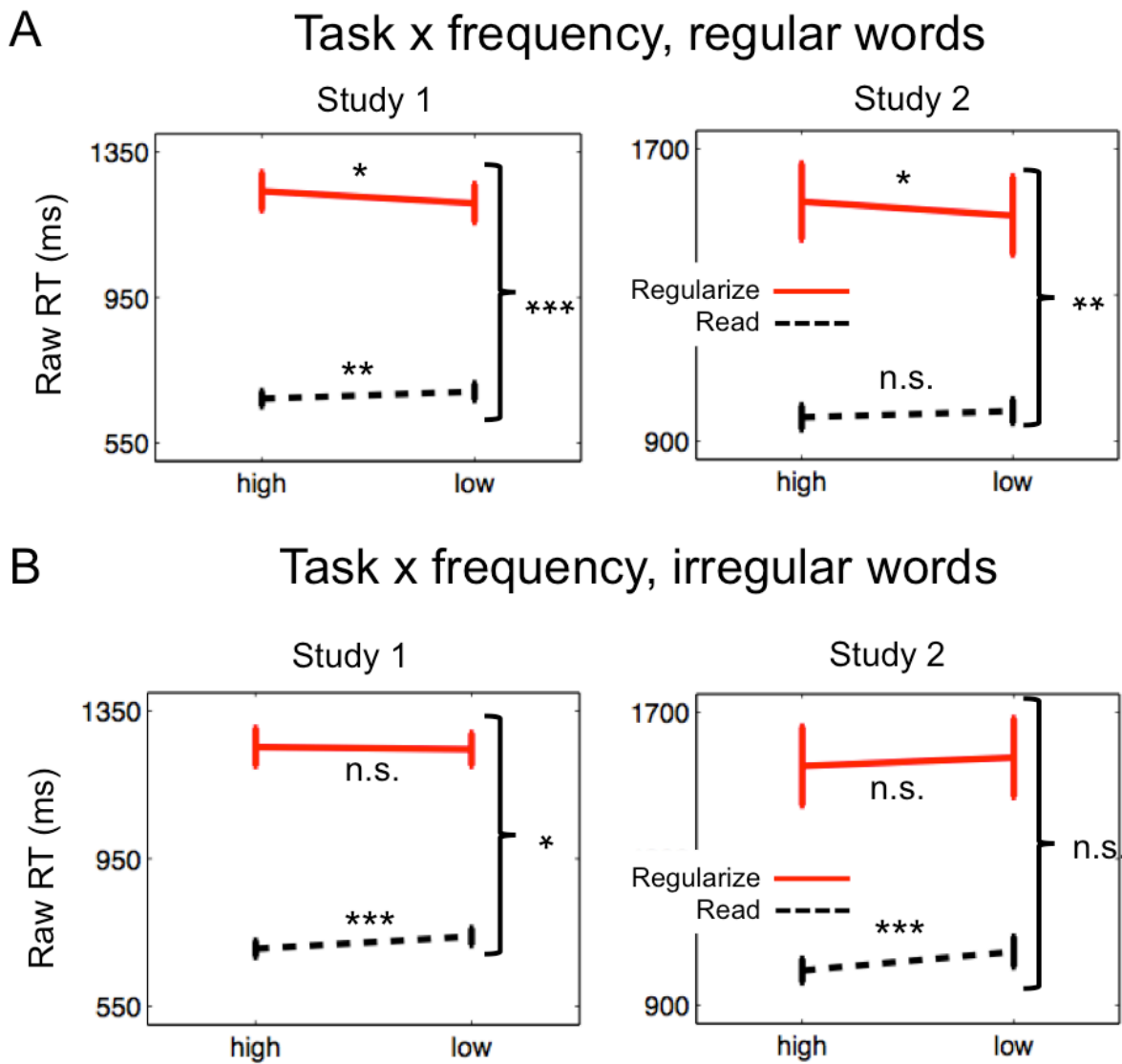
Post hoc analyses at the item-level suggested that regularity did modulate word frequency effect reversal for regularize task

The 3-way interaction of task x regularity x frequency was one of the pieces of evidence that Balota et al. cited as reason to invoke a 2-mechanism account (Balota et al. 2000). However, the present data did not yield a significant 3-way interaction, despite apparent qualitative differences between the frequency effects for the regular and irregular words (see Figure 2.4). Thus a *post hoc* analysis was conducted to investigate the possibility that treating frequency as a continuous variable (rather than a dichotomous one) may uncover a more *subtle* difference in the way that regularity modulates the influence of frequency on regularize task speed. Further supporting the appropriateness of conducting a correlation analysis of the effect of frequency is the fact that the word stimuli were selected intentionally to sample a continuous range of the log-transformed Zeno et al. Standard Frequency Index (SFI) (Zeno et al. 1995) (see Supplemental figure B.8), and the original choice to dichotomize the stimuli by a median split, while convenient, by no means represented a theoretical commitment to thinking about frequency as a simple construct of high vs. low.

In order to maximize the power to detect what was predicted to be a weak effect, item-level correlation analyses were computed using all 384 word stimuli (i.e., rather than only the short words) from Study 1 (n = 192 each of regular and irregular words). Each item was entered into two bivariate correlations, one for each task (read and regularize). The correlations between frequency and group-average zRT for each task were computed separately for regular and irregular words. As shown in Figure 2.5, frequency facilitated read zRT for both regular (R = -.176; p = .015) and irregular words (R = -.312; p < .001), as expected. Critically, frequency also significantly *inhibited* regularize task zRT for regular words (R = .242; p = .001), while the effect for irregular words is null (R = -.046; p = .524). The regular vs. irregular distinction revealed by the correlation results, while not a direct test of difference, is nonetheless consistent with Balota

et al.'s 3-way interaction of task x frequency x regularity, further motivating the exploration of the 2-mechansim hypothesis with the imaging data.

When the correlations were recomputed using only the short words from Study 1, the patterns of facilitation and inhibition remained the same. For the critical effect of inhibition of regularize reaction time for regular words, reliability was decreased to *near* statistical significance (p = .051), but the Pearson's R was very similar (R = .200), and the slope of the linear fit was comparable. Also, the effect of frequency on irregular words in the regularize task was still null when computed using the smaller stimulus set (R = -.062; p = .550).

Figure 2.5 Item-level correlations between task performance and frequency, stratified by regularity

Frequency (measured as SFI, a log-transformed measure) significantly facilitated performance on the read task for both word types (A), and inhibited performance on the regularize task for regular, but not irregular, words (B). Each word is represented in each plot, with SFI plotted on the x-axis against zRT (averaged across participants) on the y-axis. Asterisks and solid lines indicate a significant linear relationship, where ***p ≤ .001; *p ≤ .05. The dotted line indicates the absence a significant fit. Regular words are shown as orange diamonds, while irregular words are plotted as periwinkle circles. The x- and y-axis scales are the same for both plots.
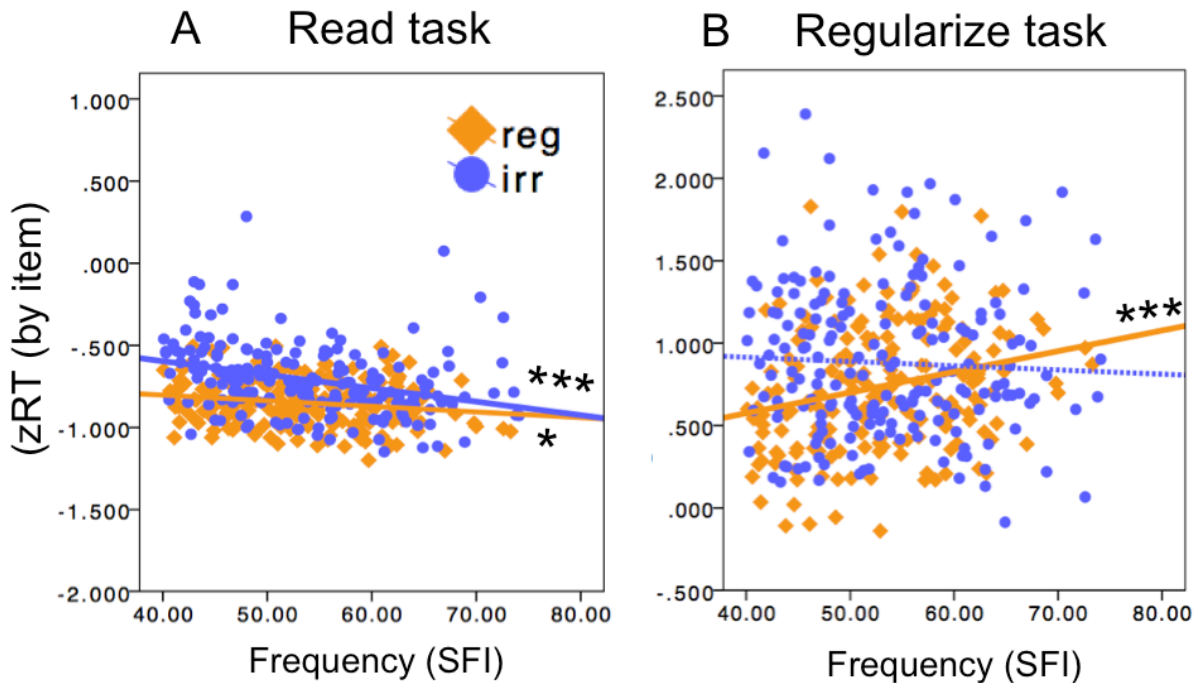
Supplemental material C.1 presents a *post hoc* correlation analysis, conducted at the subject-level, that provides further corollary support of the 2-mechanism account.

fMRI revealed substantial involvement of attentional control-related brain regions

As described in the Methods, voxelwise rmANOVAs to identify the effects of task and lexicality were prioritized for region definition (rather than effects of regularity) for two reasons: they showed the largest and most reliable behavioral effects, and they could be computed using the full stimulus set (i.e., words and pseudowords). Effects of regularity were computed on the regions defined by task and lexicality, as described below.

Task x timecourse

Many regions of the brain, in both the right and left hemispheres, showed differential activity across time that depended on task. Figure 2.6 depicts the statistical image of the interaction of task by timecourse. Although activations were widespread, the focus here is on regions involved in attentional control in relation to reading-related processing.

Figure 2.6 shows prominent involvement of regions in the Cingulo-opercular control system (purple arrows), including bilateral anterior insula/frontal operculum (aI/FO) and dorsal anterior cingulate cortex/ medial superior frontal cortex (dACC/ msFC); Frontal-parietal control system (yellow arrows), including left dorsolateral prefrontal cortex and left intraparietal sulcus (IPS); and Dorsal attention system (green arrows), including bilateral Frontal eye field (FEF). Reliable activations were also obtained in regions typically reported in studies of single word reading, including left angular gyrus (AG) and in a region in left fusiform gyrus near the putative visual word form area (pVWFA). Activations in bilateral Inferior frontal gyrus *pars opercularis* (IFG*po*) wrap under the lip of the gyrus inferiorly and are difficult to visualize in this image. Note the lack of a reliable effect in the left supramarginal gyrus (SMG)[6], consistent with our previous

---

[6] Follow-up analyses using an applied region (MNI -54, -41, +26 (Church et al. 2008) that the left SMG, despite showing a robust main effect of timecourse across read and regularize, did

work in this region showing its insensitivity to lexicality and suggesting a functional role as a

generic phonological processor (e.g., Church et al. 2008).

Figure 2.6 Statistical image of the voxelwise rmANOVA for the effect of task by timecourse
　　　　Color scale corresponds to a *z*-score range of 3.5 to 6.0. Volumetric data was displayed on an inflated cortical surface using Caret software (http://www.nitrc.org/projects/caret/; Van Essen et al. 2001); 3 transverse slices in volumetric space are also shown. Colored arrows highlight control-related regions belonging to the Cingulo-opercular control system (purple), Dorsal attention system (green) and Frontal-parietal control system (yellow).
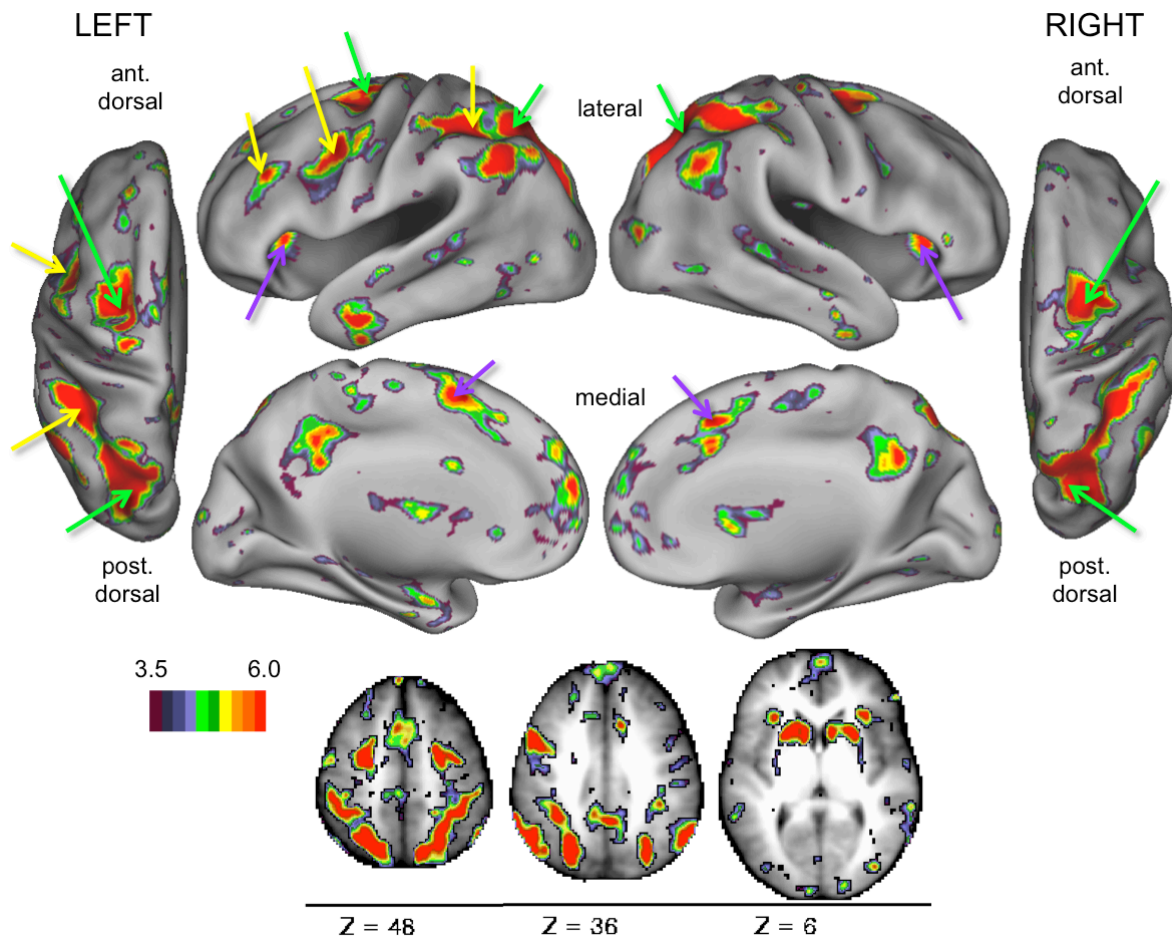
not show reliable effects of the task, lexicality x task or regularity (all sphericity corrected Z < 2.0).

Task x lexicality x timecourse

Many brain regions, spanning both hemispheres, showed significant BOLD activity changes across time that depended both on lexicality and task (see Figure 2.7). The task x lexicality x timecourse image appears similar to the task x timecourse image (compare Figures 2.6 and 2.7). Prominent activations are seen in regions belonging to the 3 control systems highlighted in the task x timecourse image in Figure 2.6: bilateral Cingulo-opercular (purple arrows); left-hemisphere Frontal-parietal (yellow arrows); and bilateral Dorsal attention (green arrows).

Region selection

The automated peak-finding algorithm identified 72 regions of interest (ROIs) in the lexicality x task x timecourse image, and 144 regions in the task x timecourse image, for a total of 216 possible ROIs. After removal of presumptive overlaps as well as regions with non-biologic timecourses (procedure described in Supplemental material A.4), 156 regions remained in the final pool (see Table 2.6 for region set).

Figure 2.7 Statistical image of the voxelwise rmANOVA for the effect of task by lexicality by timecourse

　　　Color scale corresponds to a *z*-score range of 3.5 to 6.0. Volumetric data was displayed on an inflated cortical surface using Caret software (Van Essen et al. 2001); 3 transverse slices in volumetric space are also shown. As in Figure 2.6, colored arrows highlight control-related regions belonging to the Cingulo-opercular control system (purple), Dorsal attention system (green) and Frontal-parietal control system (yellow).
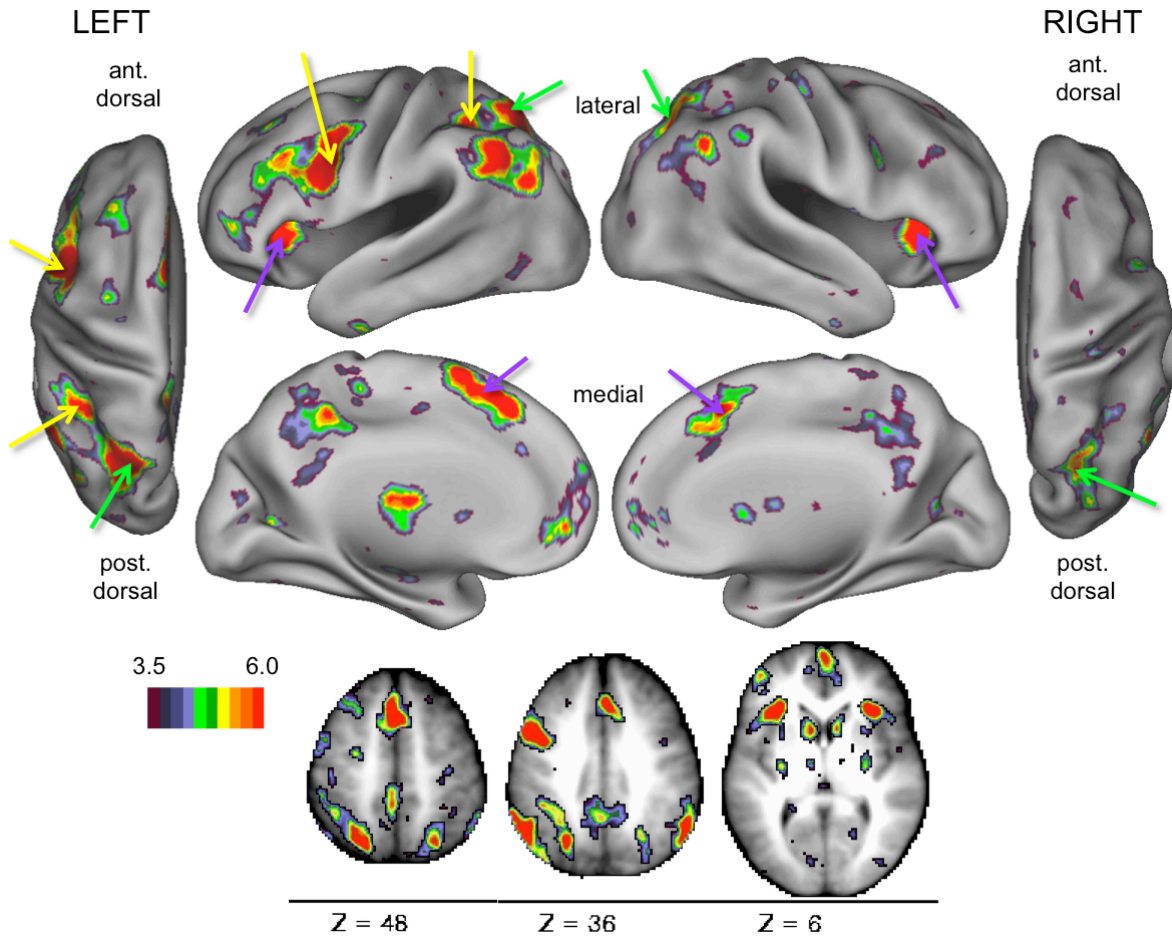
Table 2.6 List of all 156 regions' coordinates, including cluster assignment, when applicable

X[a]	Y[a]	Z[a]	Gross anatomic location	Size (cm^3)	source[b] ANOVA	Z-score from source[b] ANOVA	Dendro-gram[c]
Left hemisphere							
-23	40	43	Superior frontal gyrus	1.264	Task	4.72	A
-21	60	25	Superior frontal gyrus	0.472	Task	4.24	A
-17	64	-10	Superior frontal gyrus	0.560	Task	3.91	A
-29	-4	55	Superior frontal gyrus	1.056	LTT	4.11	CO/DA+
-27	-3	54	Superior frontal gyrus	3.968	Task	8.38	CO/DA+
-31	-22	61	Postcentral gyrus	0.752	Task	3.94	B
-52	-5	47	Precentral gyrus	1.664	Task	5.07	B
-49	-9	37	Precentral gyrus	1.464	Task	4.14	B
-24	-18	69	Precentral gyrus	1.440	Task	5.07	B
-48	9	32	Middle frontal gyrus	4.000	LTT	7.81	LFP
-28	28	46	Middle frontal gyrus	2.448	LTT	4.82	A
-12	-5	70	Medial frontal gyrus	1.016	Task	4.19	B
-11	53	8	Medial frontal gyrus	1.000	LTT	3.97	A
-31	25	1	Anterior insula	2.272	LTT	7.20	CO/DA+
-42	21	-2	Frontal operculum	2.240	LTT	6.60	B
-44	34	22	Prefrontal cortex	1.504	Task	5.62	LFP
-46	38	8	Inferior frontal gyrus	1.048	LTT	4.21	LFP
-45	45	-2	Inferior frontal gyrus	1.368	LTT	4.88	B
-49	21	-3	Inferior frontal gyrus	0.400	Task	3.67	LFP
-58	-63	29	Inferior parietal lobe	2.408	Task	7.37	A
-50	-36	49	Inferior parietal lobe	1.064	LTT	4.19	LFP
-45	-70	44	Inferior parietal lobe	1.952	Task	6.19	A
-22	-39	55	Inferior parietal lobe	0.696	Task	3.80	B
-41	-43	45	Intraparietal sulcus	2.376	LTT	5.67	LFP
-41	-39	45	Intraparietal sulcus	2.944	Task	7.36	LFP
-34	-47	38	Intraparietal sulcus	2.136	Task	6.12	LFP
-62	-50	38	Angular gyrus	1.856	LTT	5.57	A

-54	-61	41	Angular gyrus	2.456	Task	8.20	A
-47	-71	32	Angular gyrus	3.160	LTT	5.88	A
-21	-65	5	Lingual gyrus	0.400	LTT	4.09	B
-23	-68	44	Precuneus	3.752	Task	8.98	LFP
-37	-85	7	Middle occipital gyrus	0.560	Task	3.94	B
-28	-87	18	Middle occipital gyrus	0.240	LTT	3.54	B
-25	-76	24	Middle occipital gyrus	0.712	LTT	3.73	CO/DA+
-49	-60	-9	Fusiform gyrus	1.656	Task	4.47	LFP
-43	-62	-6	Fusiform gyrus	1.400	LTT	4.44	B
-57	-10	16	Superior temporal gyrus	0.800	Task	3.64	B
-64	-47	-3	Middle temporal gyrus	0.248	Task	3.68	A
-59	-11	-11	Middle temporal gyrus	0.664	Task	4.03	A
-55	-33	-3	Middle temporal gyrus	0.608	Task	4.08	A
-44	-15	19	Lateral sulcus	0.704	Task	4.08	B
-41	-68	1	Inferior temporal gyrus	1.064	Task	5.08	B
-29	-9	0	Putamen	0.872	LTT	4.48	B
-24	13	0	Putamen	2.240	Task	7.04	B
-13	10	8	Caudate	3.088	Task	8.14	B
-28	-27	-11	Hippocampus	0.352	Task	3.54	A
-27	3	-37	Uncus	0.840	Task	4.28	A
-30	-63	-21	Cerebellum	0.408	LTT	3.59	B
-22	-75	-20	Cerebellum	1.024	Task	4.64	B
-21	-73	-22	Cerebellum	0.712	LTT	4.10	B
-17	-42	-17	Cerebellum	0.384	Task	3.53	A
Interhemisphere							
-4	27	-6	Anterior cingulate / medial frontal gyrus	0.576	Task	4.08	A
-1	20	49	Anterior cingulate / medial frontal gyrus	3.840	LTT	8.16	CO/DA+
2	29	37	Anterior cingulate /medial frontal gyrus	3.032	LTT	7.12	CO/DA+
-1	11	56	Anterior cingulate /medial frontal gyrus	3.280	Task	6.94	CO/DA+
-2	8	26	Anterior cingulate gyrus	0.272	Task	3.91	CO/DA+

-5	-52	22	Posterior cingulate gyrus	0.656	LTT	3.85	A
-3	-37	45	Posterior cingulate gyrus	2.464	LTT	5.70	A
8	-52	22	Posterior cingulate gyrus	0.616	LTT	3.88	A
8	-31	43	Posterior cingulate gyrus	1.056	LTT	4.28	A
-10	50	24	Medial frontal gyrus	1.112	Task	4.40	A
-2	62	14	Medial frontal gyrus	3.072	Task	5.87	A
-1	52	-11	Medial frontal gyrus	1.720	Task	5.16	A
-1	58	32	Medial frontal gyrus	2.440	Task	5.04	A
1	60	-2	Medial frontal gyrus	2.592	LTT	5.80	A
3	-5	51	Medial frontal gyrus	1.152	Task	4.11	B
4	42	-8	Medial frontal gyrus	0.672	Task	4.07	A
8	49	-2	Medial frontal gyrus	1.088	LTT	4.26	A
10	20	34	Medial frontal gyrus	1.232	Task	5.22	CO/DA+
-10	26	34	Medial frontal gyrus	0.512	Task	3.69	CO/DA+
-3	51	45	Medial frontal gyrus	0.296	Task	3.80	A
-8	-30	52	Paracentral lobule	2.056	Task	4.77	A
-3	-38	65	Paracentral lobule	1.336	LTT	3.92	A
1	-17	59	Paracentral lobule	1.256	Task	4.35	B
-9	-100	12	Cuneus	1.056	Task	4.28	A
-5	-81	12	Cuneus	0.720	LTT	3.82	B
10	-94	28	Cuneus	0.416	Task	4.08	A
-3	-77	6	Lingual gyrus	0.320	Task	3.77	B
-7	-53	56	Precuneus	1.104	LTT	4.00	A
-5	-44	39	Precuneus	2.736	Task	5.66	A
4	-49	36	Precuneus	2.536	Task	6.35	A
6	-51	57	Precuneus	0.488	LTT	3.70	A
10	-60	35	Precuneus	1.448	Task	4.87	A
10	-46	41	Precuneus	2.040	LTT	4.35	A
-6	-13	12	Thalamus	2.536	LTT	5.82	B
-1	-8	13	Thalamus	1.216	Task	5.86	B
0	-20	12	Thalamus	1.320	Task	6.37	B
-10	11	4	Caudate	1.832	LTT	5.44	B
-8	-59	-11	Cerebellum	0.792	Task	4.31	B

-2	-46	-7	Cerebellum	0.848	Task	4.34	A
7	-51	-16	Cerebellum	0.480	Task	3.78	A
7	-40	-10	Cerebellum	0.552	Task	3.86	A
9	-66	-14	Cerebellum	0.720	Task	3.93	B
Right hemisphere							
20	55	35	Superior frontal gyrus	0.600	Task	3.89	A
24	31	50	Superior frontal gyrus	0.496	Task	4.06	A
36	-4	48	Superior frontal gyrus	2.304	Task	5.87	B
27	-4	56	Superior frontal gyrus	2.896	Task	7.33	B
43	-17	62	Postcentral gyrus	0.512	LTT	3.70	B
30	-31	66	Postcentral gyrus	0.552	Task	3.56	B
36	-22	55	Precentral gyrus	1.368	LTT	4.14	B
43	-16	55	Precentral gyrus	0.968	Task	4.48	B
52	-8	37	Precentral gyrus	0.784	Task	3.98	B
31	-18	62	Precentral gyrus	1.480	Task	5.12	B
45	3	37	Middle frontal gyrus	0.560	Task	4.35	CO/DA+
53	12	36	Middle frontal gyrus	0.376	Task	3.72	CO/DA+
45	21	1	Frontal operculum	1.376	LTT	4.30	CO/DA+
45	21	-2	Inferior frontal gyrus	0.424	Task	3.85	B
45	33	28	Inferior frontal gyrus	0.376	LTT	3.63	B
47	9	25	Premotor cortex	0.920	LTT	4.70	CO/DA+
11	51	6	Medial frontal gyrus	0.712	Task	3.91	A
14	-19	60	Medial frontal gyrus	0.928	Task	4.28	B
14	1	58	Medial frontal gyrus	1.720	Task	5.08	B
34	25	0	Anterior insula	2.568	LTT	7.12	CO/DA+
19	-29	54	Paracentral lobule	0.736	Task	4.41	B
12	-19	72	Paracentral lobule	0.264	Task	3.57	B
24	-62	55	Superior parietal lobe	2.232	Task	8.22	CO/DA+
29	-69	39	Superior parietal lobe	3.128	Task	8.41	CO/DA+
29	-52	54	Superior parietal lobe	2.848	Task	7.13	CO/DA+
35	-42	51	Superior parietal lobe	1.216	LTT	3.94	CO/DA+
40	-38	48	Superior parietal lobe	3.728	Task	7.51	B

53	-24	46	Superior parietal lobe	1.192	Task	4.94	B
40	-26	33	Inferior parietal lobe	0.816	Task	5.16	B
50	-50	29	Inferior parietal lobe	1.176	LTT	4.45	A
50	-34	32	Inferior parietal lobe	0.432	LTT	3.68	A
52	-66	37	Inferior parietal lobe	1.912	LTT	5.46	A
55	-49	46	Inferior parietal lobe	1.288	LTT	4.67	A
56	-57	40	Inferior parietal lobe	2.824	Task	7.08	A
57	-61	25	Inferior parietal lobe	1.744	Task	5.69	A
64	-37	40	Inferior parietal lobe	0.672	Task	3.75	A
56	18	-10	Superior temporal sulcus	0.200	Task	3.56	B
63	-7	2	Superior temporal gyrus	0.264	Task	3.69	B
38	-24	11	Transverse temp. gyrus	0.344	Task	3.97	B
35	-19	21	Lateral sulcus	0.336	Task	4.11	B
62	-59	3	Middle temporal gyrus	0.456	Task	4.14	A
50	-35	1	Middle temporal gyrus	0.960	Task	4.44	B
51	-24	-7	Middle temporal gyrus	0.496	Task	3.77	B
63	-49	7	Middle temporal gyrus	0.808	Task	4.31	A
60	-9	-18	Medial temporal gyrus	0.312	LTT	3.87	A
30	-73	26	Middle occipital gyrus	1.168	LTT	4.66	B
39	-81	10	Middle occipital gyrus	1.248	Task	5.05	B
49	-60	-6	Fusiform gyrus	0.600	Task	3.82	CO/DA+
15	-70	54	Precuneus	2.304	Task	7.21	B
20	-62	10	Precuneus	0.488	LTT	4.10	B
11	-16	13	Thalamus	0.552	LTT	3.83	B
17	-15	-13	Brainstem	0.288	Task	3.75	A
25	12	0	Putamen	1.768	Task	6.80	B
30	-11	-1	Putamen	1.584	LTT	4.65	B
30	10	0	Putamen	0.432	LTT	3.91	B
11	7	8	Caudate	1.320	LTT	5.06	B
12	11	8	Caudate	2.152	Task	7.68	B
14	4	17	Caudate	1.672	Task	5.21	B
18	-26	17	Caudate	0.320	Task	3.73	B
24	-8	-20	Amygdala/ hippocampus	0.688	Task	4.31	A

14	-77	-21	Cerebellum	0.984	LTT	4.18	B
28	-69	-24	Cerebellum	1.216	LTT	4.47	B
17	-95	11	Cuneus	0.736	Task	5.09	B

Note. [a]XYZ coordinates are reported in MNI space. [b]For source ANOVA, LTT = lexicality x task x timecourse; Task = task x timecourse. The Z-score from the source ANOVA is from the highest-order interaction from the voxelwise source ANOVA i.e., (task x timecourse) or (lexicality x task x timecourse). [c]For dendrogram column, members of the CO/DA+ and LFP clusters are marked as such; all other regions are labeled as to their general location on the dendrogram, where A = left side of tree (mostly negative timecourses) and B = right side of tree (mostly positive timecourses) (see Figure 2.8).

Hierarchical clustering revealed 2 dissociable control-related clusters

BOLD hemodynamic response function (HRF) shapes carry information about the processing of particular brain regions that complements the quantification provided by statistical testing of effects of interest. In particular, it is possible to identify sets of brain regions that are functionally related for a particular task by querying for similarity in the regions' BOLD timecourse profiles (e.g., Ploran et al. 2007; Church et al. 2011).

UPGMA clustering of the final set of 156 regions produced a dendrogram with relatively high fidelity to the underlying structure (cophenetic R = 0.82; see Figure 2.8). Consistent with the utility of the clustering technique for identifying important inter-region relationships, several plausible functional neuroanatomical distinctions were recognizable in the tree, including, for example, the clustering of classically described Default Mode Network regions (see blue and green clusters on the far left of the dendrogram). The objective here, however, was to identify clusters containing attentional control regions. At a similarity threshold of (1-r) = 0.15, two candidate clusters were identified, each composed primarily of control- and reading-related regions (see Figure 2.9).

Figure 2.8 Dendrogram resulting from the hierarchical clustering of 156 regions computed using a UPGMA method in Matlab

Highlighted are the 2 clusters identified as being comprised chiefly of control-related regions: a Cingulo-opercular/ Dorsal attention + (CO/DA+) cluster (purple), including 20 regions, and a Left Frontal-parietal (LFP) cluster (yellow), including 10 regions. Red horizontal line indicates clustering threshold of (1-r) = 0.15. A and B mark the left and right sides of the dendrogram (see Table 2.6). Cophenetic R = 0.82.
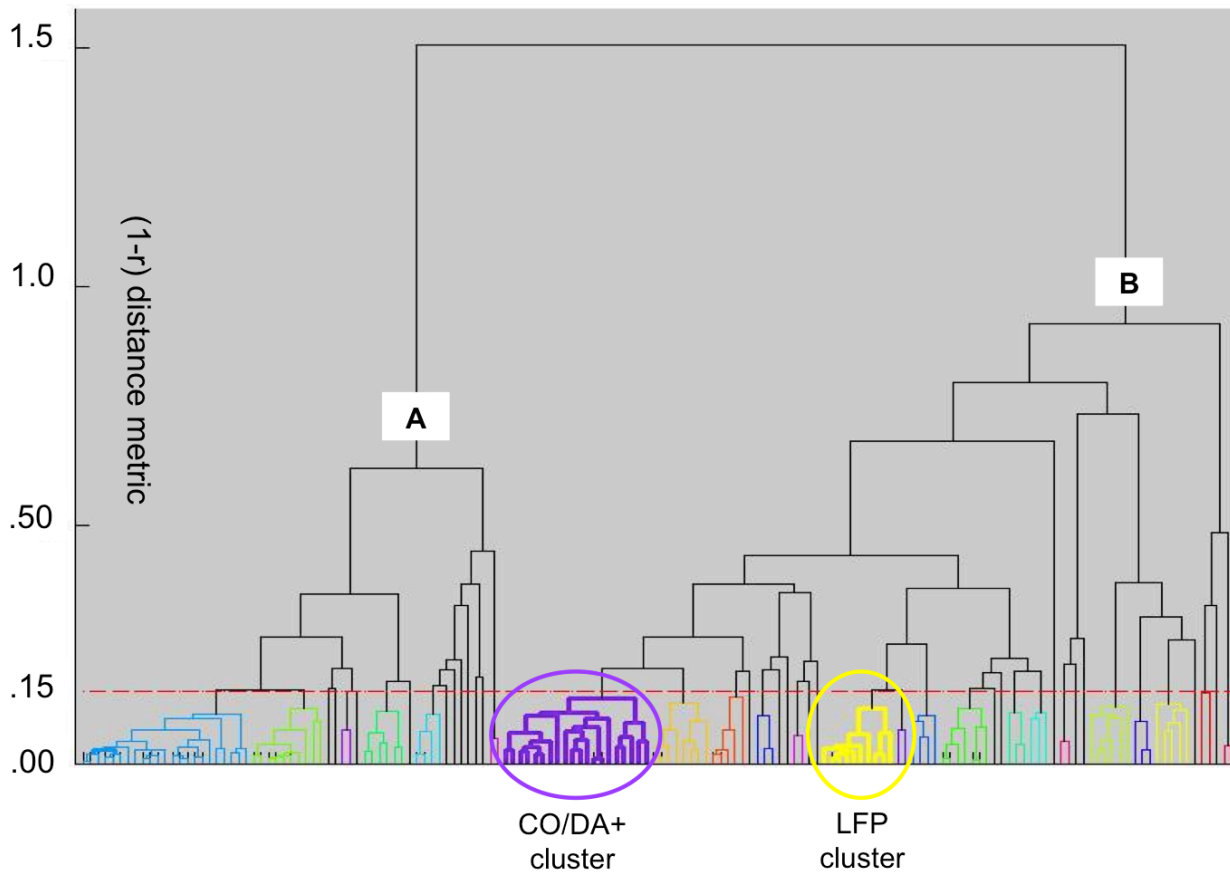
The cluster shown in purple in Figure 2.9 includes several regions of the Cingulo-

opercular control system bilaterally, as well as a handful of mostly (but not entirely) right-

lateralized Dorsal attention system regions (referred to as CO/DA+; see Supplemental figure

C.2 for the image including the network borders). The cluster projected in yellow in Figure 2.9 is

entirely left-lateralized, and includes almost exclusively regions that fall into the Frontal-parietal

control network (abbreviated LFP; again, see Supplemental figure C.2).

Figure 2.9 Two candidate attentional control systems for reading, identified using hierarchical clustering
Surface projection of the Cingulo-opercular/ Dorsal attention plus (CO/DA+) cluster, shown in purple (n = 20 regions), and the left Frontal-parietal (LFP) cluster, shown in yellow (n = 10 regions), defined from the clustering depicted in Figure 2.8. Clusters projected using Caret software (Van Essen et al. 2001). See also Supplemental figure C.2 for a version of this figure that includes network areal borders.

76
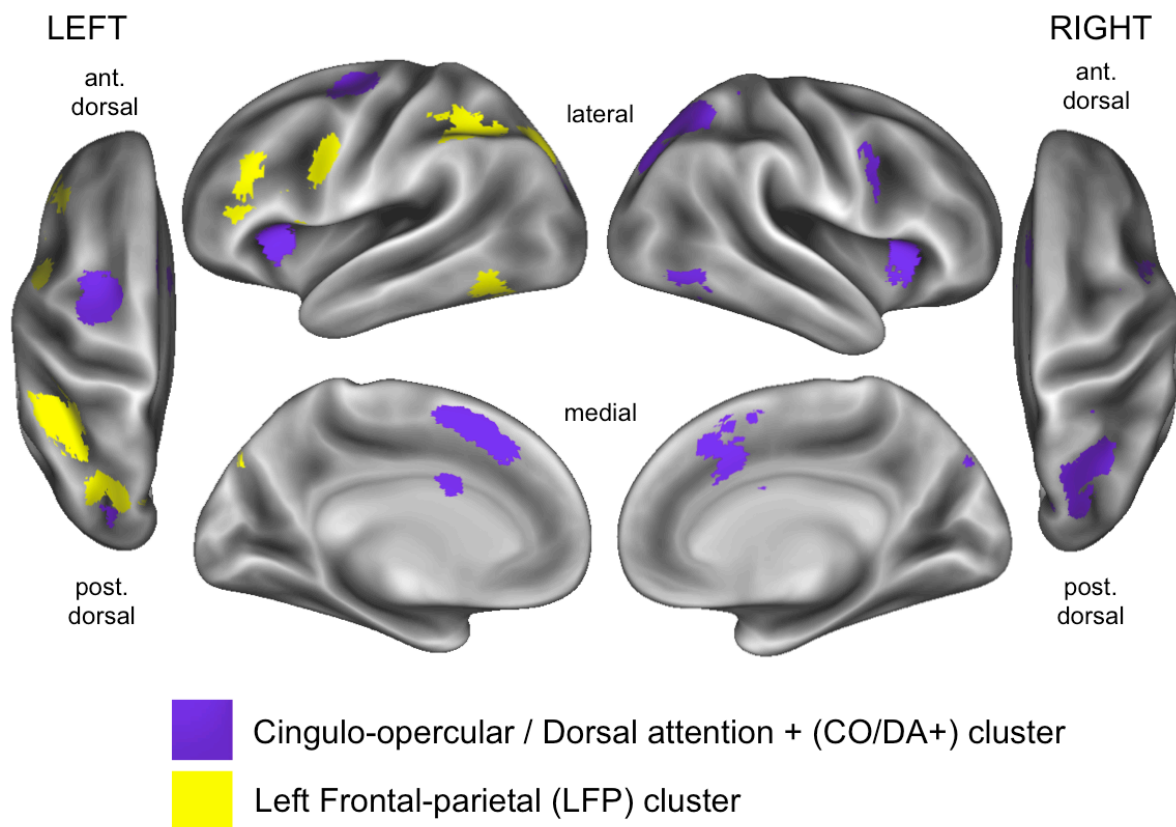
Statistical tests of differences between the clusters

Omnibus rmANOVAs and planned *post hocs* were computed in order to test the stated predictions of the 2-mechanism model (outlined in Table 2.4) and determine which of various task- and stimulus- combinations distinguished the 2 cluster sets. Figures 2.10-2.15 depict the timecourses, averaged across the 2 clusters separately, for the task/stimulus combinations of interest, as well as the results of the rmANOVAs and planned *post hocs*. For the analyses, each region from the CO/DA+ and LFP clusters was treated as a "subject," and cluster assignment was treated as the grouping variable. Note that the timecourses depict all 8 timepoints in the timecourse, even though only the first 6 frames were used to compute the clustering (see explanation above). The *post hoc* analyses likewise used all 8 timepoints of the timecourses.

First, the 2 omnibus rmANOVAs confirmed that cluster membership reliably modulated both the interaction of lexicality x task x timecourse ($p = .004$, Greenhouse-Geisser corrected; see Figure 2.10) and the interaction of regularity x task x timecourse ($p \leq .001$; see Figure 2.13). We describe first the interactions of lexicality, task, timecourse and cluster, then the interactions of regularity, task, timecourse and cluster.

Sensitivity to lexicality

Each individual yellow and purple timecourse, indicating a cluster's response to a single task/lexicality condition, was significantly different from zero. Furthermore, for all 4 combinations of lexicality and task, the LFP cluster showed a significantly larger % BOLD signal change, across timecourse, than the CO/DA+ cluster (all $p \leq .001$; see Figure 2.10). Because all 4 conditions showed a significant effect of cluster on the timecourse (and each timecourse showed a significant main effect of time), the effects of lexicality and cluster were next examined within each task separately, as well as the effects of task and cluster for each lexicality class separately.

Figure 2.10 Lexicality x task x timecourse x cluster

Average BOLD timecourses for each of the 2 control clusters, for each of 4 combinations of lexicality and task. Each MR frame was 3 s. Asterisks indicate a significant interaction (rmANOVA: *** $p \leq .001$). For all 4 conditions, the LFP cluster (yellow) showed a significantly larger % BOLD signal change, across time, than the CO/DA+ cluster (light purple). The 4-way interaction of task x lexicality x timecourse x cluster was also significant, as indicated in the middle of the plot, as was each main effect of timecourse.
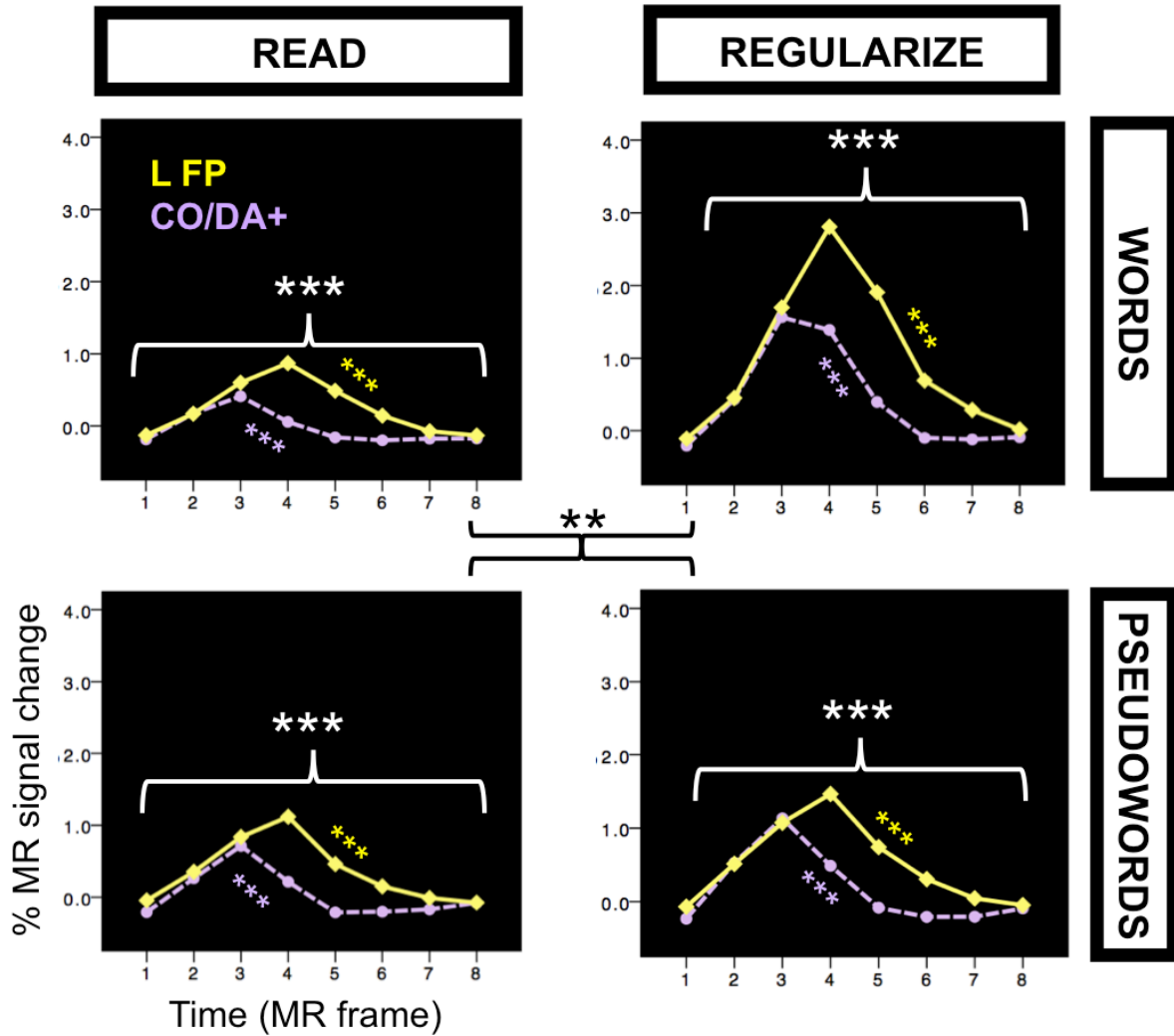
The first of the 2 planned rmANOVA pairs is depicted in Figure 2.11. rmANOVAs conducted separately for each task revealed that although the lexicality effect for the read task was significant in each cluster (both $p \leq$ .001; top panel), the effect was larger in the CO/DA+ regions than the LFP regions (p = .035). The bottom panel shows that the (reversed) lexicality effect was, on the other hand, significantly larger in the LFP regions than the CO/DA+ regions ($p \leq$.005), although lexicality influenced both region sets significantly (both $p \leq$ .001).

Analyses conducted separately for words and pseudowords, depicted in Figure 2.12, were likewise revealing. While both region sets showed reliable effects of task for words (both $p \leq$ .001; top panel), the task effect for words was greater in the LFP regions than the CO/DA+ regions ($p \leq$ .001). The pattern was different for pseudowords (bottom panel). The CO/DA+ regions ($p \leq$ .001), but not the LFP regions (p = .129), showed a significant effect of task across the timecourse for pseudowords. However, the effects were small, and the interaction of task x timecourse x cluster for pseudowords did not reach significance (p = .160).

The timecourses of the individual regions of both clusters generally recapitulated the patterns of task and lexicality described at the group level (i.e., the group average timecourses were not driven by outliers within each cluster).

Figure 2.11 Lexicality x timecourse x cluster, each task separately

Average BOLD timecourses for each of the 2 control clusters, for each of 4 combinations of lexicality and task. The top panel shows that the lexicality effect for the read task was significantly larger in the CO/DA+ regions than the LFP regions (i.e., a significant interaction of lexicality x timecourse x cluster, for read). The bottom panel shows that the (reversed) lexicality effect for the regularize task was significantly larger in the LFP regions than the CO/DA+ regions (i.e., a significant interaction of lexicality x timecourse x cluster, for regularize). The ovals highlight the larger peak lexicality difference for each task. Asterisks indicate significant interactions (rmANOVA: * $p < .05$; ** $p \leq .005$, *** $p \leq .001$).
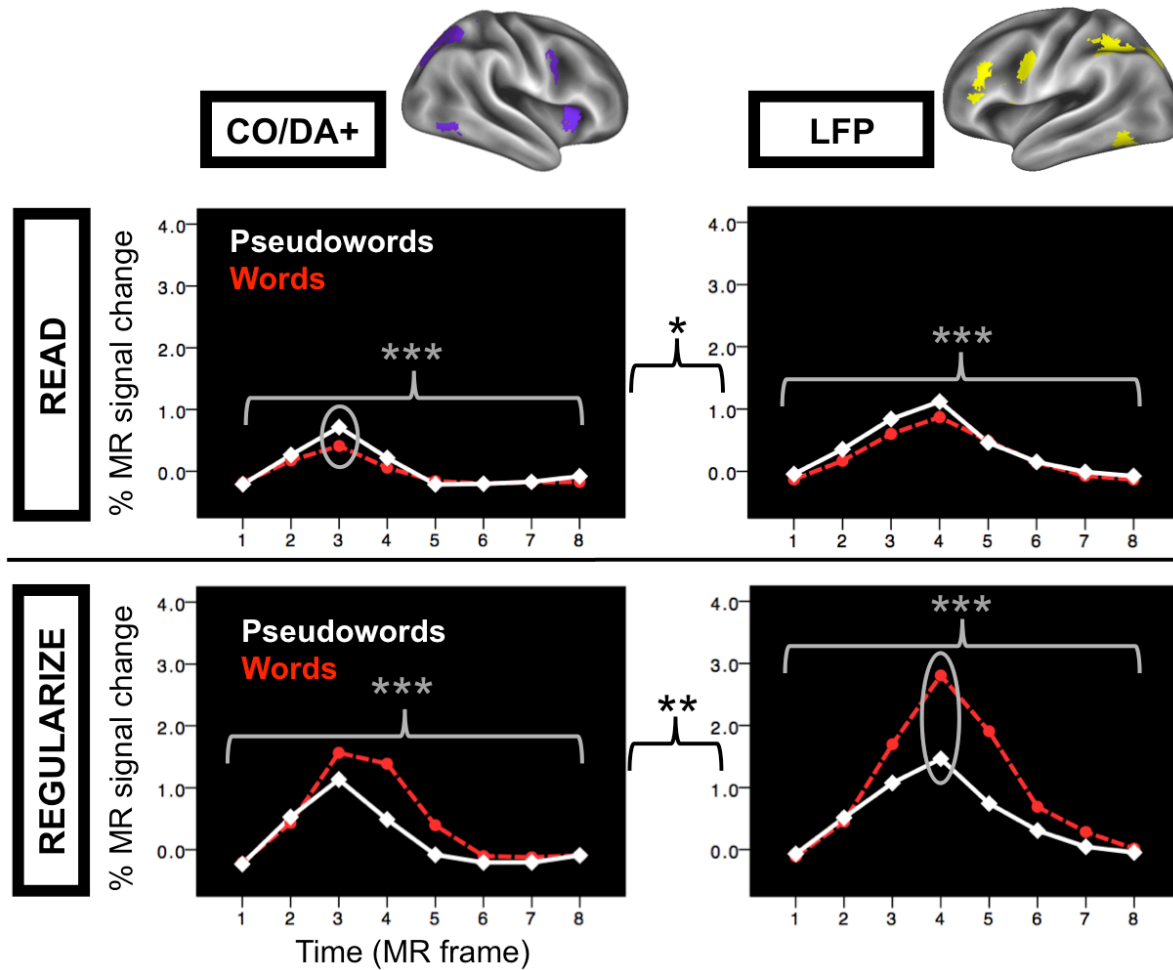
Figure 2.12 Task x timecourse x cluster, words and pseudowords separately

Average BOLD timecourses for each of the 2 control clusters, for each of 4 combinations of lexicality and task; different timecourse relationships are emphasized here as compared to Figure 2.11. The top panel shows that the task effect for words was significantly larger in the LFP regions than the CO/DA+ regions (i.e., a significant interaction of task x timecourse x cluster, for words). The bottom panel shows that the 2 sets of regions did *not* significantly differ with regard to the size of the task effect for pseudowords (i.e., a null interaction of task x timecourse x cluster, for pseudowords; *p* = 0.16). The oval highlights the larger peak task difference for the words. Asterisks indicate significant interactions (rmANOVA: *** $p \leq .001$).
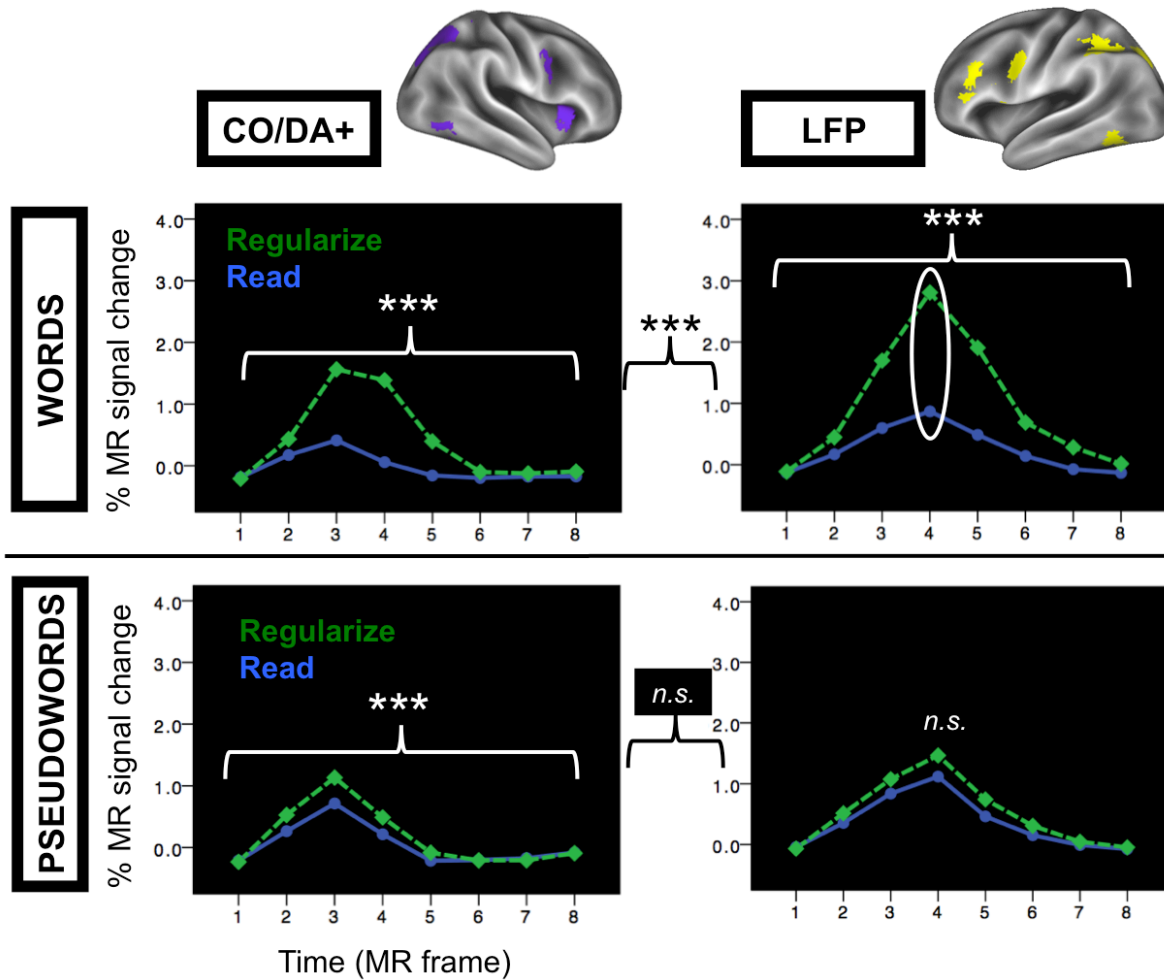
Sensitivity to regularity

Like Figure 2.10, Figure 2.13 shows timecourses, averaged across each of the 2 control clusters, for 4 conditions, but now those conditions are regularity (2) x task (2). Each individual yellow and purple timecourse, indicating a cluster's response to a single condition, was significantly different from zero. Furthermore, for all 4 conditions except for regular words in the read task, the LFP cluster's BOLD response was significantly larger, across time, than the CO/DA+ cluster's response (for regular words, read, p = .106; other 3 conditions, $p \leq .001$). Since the clusters both show a positive and not significantly different response to regular words in the read task, one of the other 3 task/regularity combinations depicted in Figure 2.13 likely drove the significant omnibus interaction.

Figure 2.13 Regularity x task x timecourse x cluster

Figure is similar to Figure 2.10, except it shows average timecourses for the 2 control clusters, for 4 conditions of regularity and task. Asterisks indicate a significant interaction (rmANOVA: * p < .05; *** p ≤ .001). For all 4 conditions except for regular words in the read task, the LFP cluster (yellow) showed a significantly larger % BOLD signal change, across time, than the CO/DA+ cluster (light purple). The 4-way interaction of task x regularity x timecourse x cluster was also significant, as indicated, as was each main effect of timecourse.
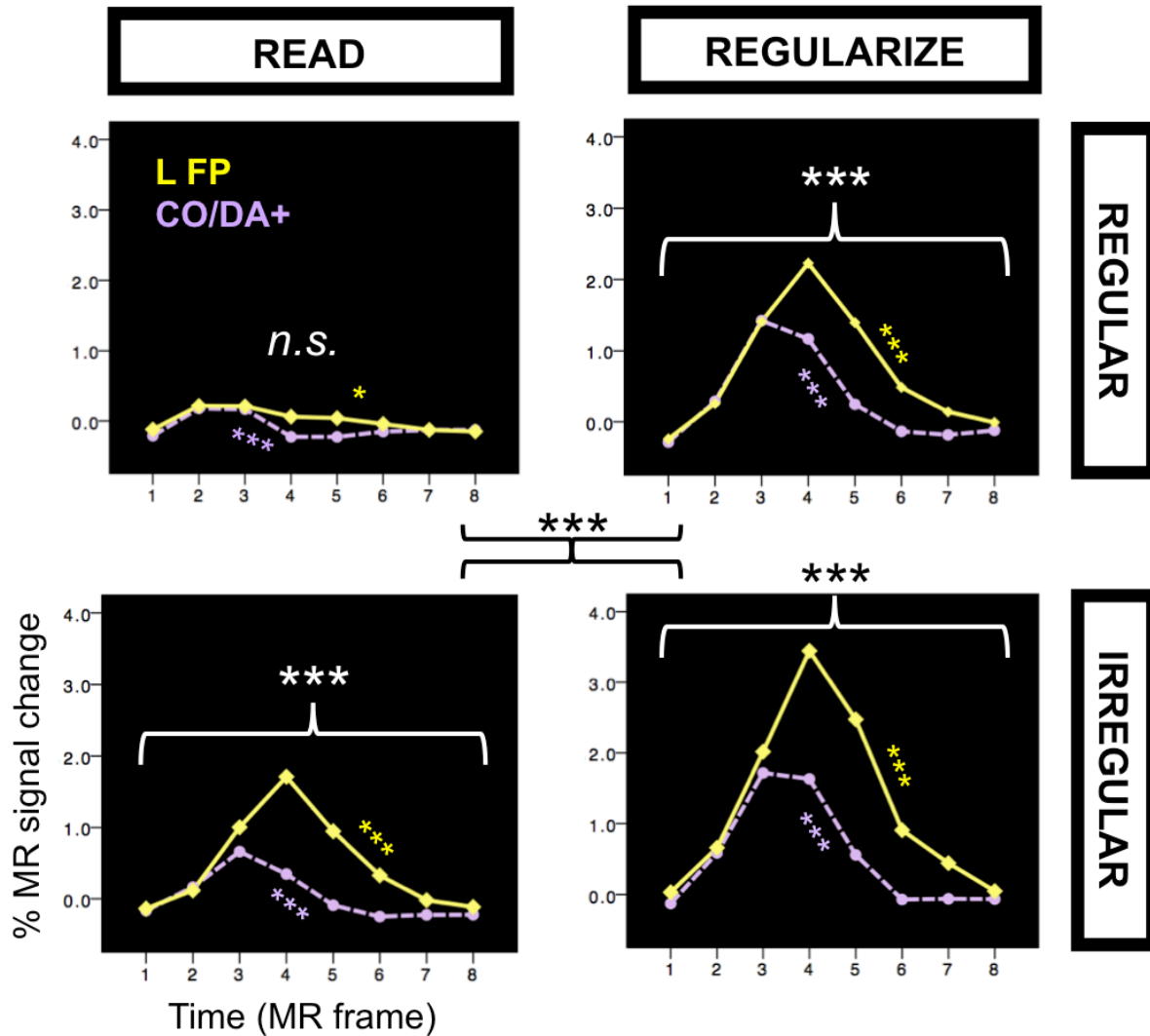
Figures 2.14 and 2.15 depict the timecourses underlying the 2 planned rmANOVA pairs examining regularity for each task, and task for each class of regularity. The LFP cluster was shown statistically to be more sensitive to regularity than the CO/DA+ cluster. Figure 2.14 shows that for both tasks considered separately, the effect of regularity (irregular > regular) was more pronounced in the LFP cluster than the CO/DA+ cluster (all $p \leq .001$).

Figure 2.14 Regularity x timecourse x cluster, each task separately
Average BOLD timecourses for each of the 2 control clusters, for each of 4 combinations of regularity and task. For both tasks (i.e., top and bottom panels), the regularity effect was significantly larger in the LFP regions than the CO/DA+ regions (i.e., a significant interaction of regularity x timecourse x cluster, for each task separately). The ovals highlight the larger peak regularity difference for each task. Asterisks indicate a significant interaction (rmANOVA: *** $p \leq .001$).
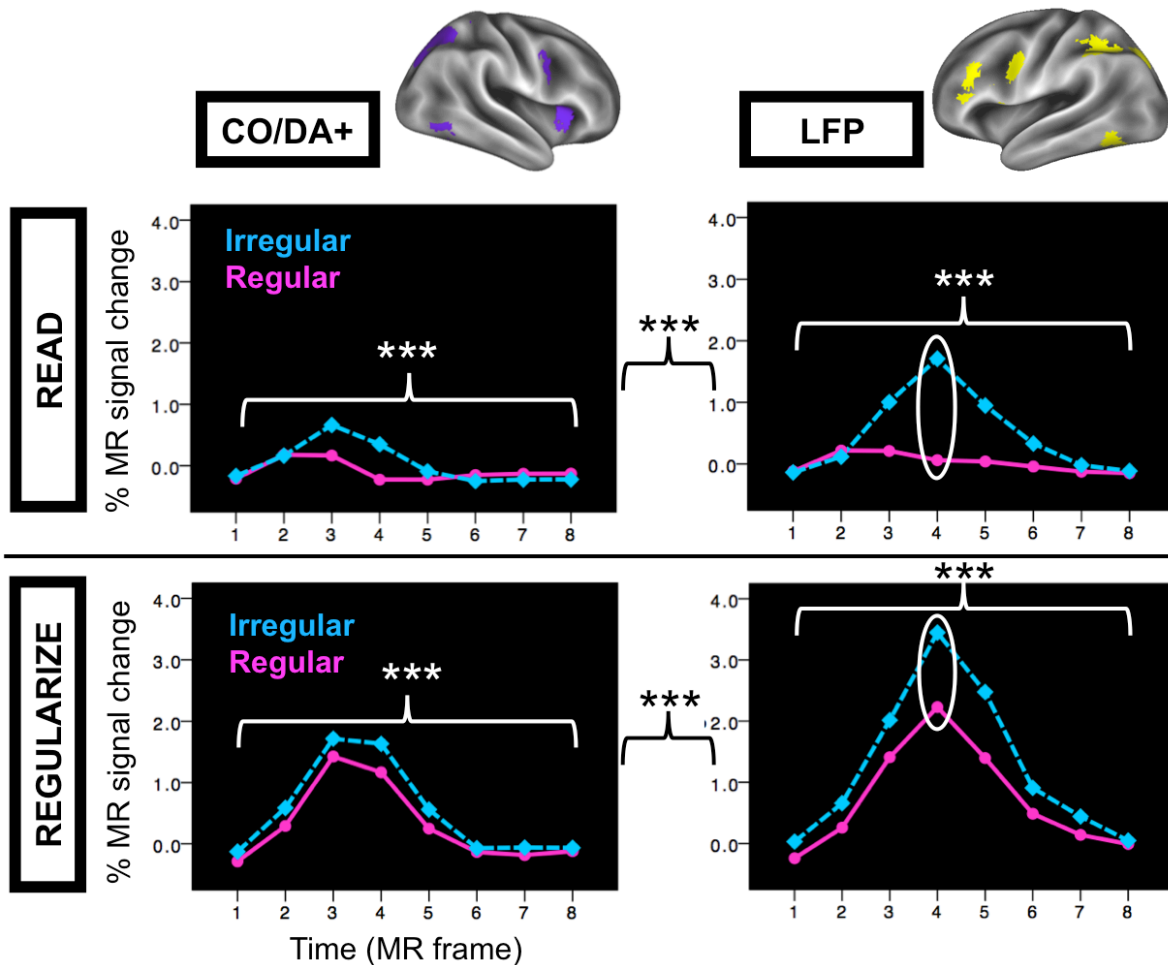
Likewise, Figure 2.15 illustrates that for both classes of regularity considered separately, the effect of task was more pronounced in the LFP cluster than the CO/DA+ cluster (all p ≤ .001).

The timecourses of the individual regions of both clusters generally recapitulated the patterns of task and regularity described at the group level (i.e., the group average timecourses were not driven by outliers within each cluster).

Figure 2.15 Regularity x task x cluster, regular and irregular words separately
Average BOLD timecourses for each of the 2 control clusters, for each of 4 combinations of regularity and task. Different timecourse relationships are emphasized here as compared to Figure 2.14. For both classes of regularity (i.e., top and bottom panels), the task effect was significantly larger in the LFP regions than the CO/DA+ regions (i.e., a significant interaction of task x timecourse x cluster, for regular and irregular words separately). The ovals highlight the larger peak task difference for each of regular and irregular words. Asterisks indicate a significant interaction (rmANOVA: *** p ≤ .001).
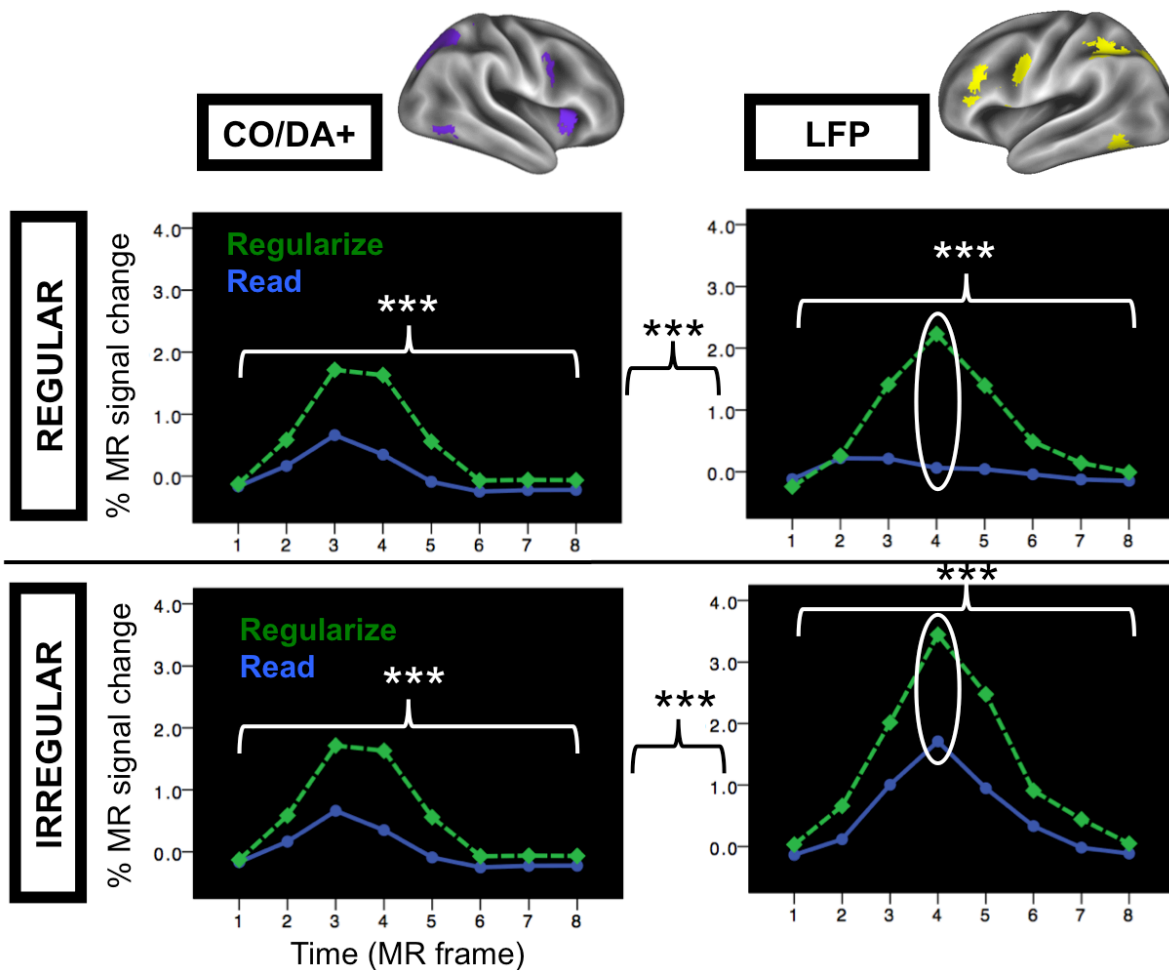
Magnitude, timing and HRF shape differences between the clusters

In addition to the information about specific statistical distinctions, 3 general featural differences between the clusters' average timecourses are apparent in Figures 2.10-2.15.

First, the overall magnitude of % BOLD signal change across time (24 s) tended to be lower in the CO/DA+ cluster than the LFP cluster.

Secondly, in terms of time-to-peak, the CO/DA+ regions tended to peak at timepoint 2 or 3 for both tasks, whereas the LFP regions almost uniformly peaked later, at timepoint 4. A Fisher's Exact Test of Independence confirmed that the CO/DA+ and LFP clusters differed significantly from one another with regard to the number of regions that peaked at timepoint 4 in both the read task (one-tailed $p \leq .001$; 0/20 for CO/DA+ vs. 8/10 for LFP) and, separately, in the regularize task (one-tailed $p \leq .001$; 1/20 for CO/DA+ vs. 10/10 for LFP).

Perhaps most informatively, the shapes of the timecourses differed notably between the 2 clusters. Specifically, not only did the CO/DA+ regions peak earlier than the LFP regions, but the BOLD responses of the CO/DA+ regions also returned to baseline more rapidly than did the LFP regions, i.e., the CO/DA+ regions showed relatively narrow HRFs, while the LFP regions showed relatively wide, extended HRFs.

Cluster "neighbors" for CO/DA+ vs. LFP

Although the analyses described here have focused on the 2 control region clusters of interest, many other brain regions were identified in the whole-brain ANOVAs and are therefore represented in the dendrogram. Because proximity in cluster space suggests similarity in terms of function, it is useful to know which regions sit most proximally to the clusters of interest in the dendrogram. CO/DA+ regions were nearest on the dendrogram to bilateral middle occipital (visual) regions and additional Dorsal attention regions, including medial superior frontal regions. LFP regions were, instead, closest to regions including parts of bilateral inferior frontal gyrus, *pars opercularis* (IFGpo), bilateral cerebellum and bilateral mouth somatomotor, as well as right

superior temporal sulcus. See Supplemental figures C.3 and C.4 for more detailed renderings of the clustering relationships within each region set, as well as the anatomical locations of the closest dendrogram neighbors of each.

Subsidiary imaging analyses

Several analyses were conducted to ensure that the observed imaging effects were not driven by potential confounds. See Supplemental materials A.3a-c for information about reaction time regression, motion censoring, and an alternate preprocessing method.

Discussion

Fluent reading involves the complex coordination of processes including sensory, linguistic and articulatory-motor; the role of attentional control in driving fluent reading to the point of automaticity (to the extent that automaticity is achieved) remains incompletely understood. Relatedly, it remains unclear how attentional control can be strategically allocated to particular aspects of reading in a skilled adult reader, a better understanding of which may illuminate the more general question of whether a tradeoff exists between the achievement of expertise vs. the ability to perform the various components of a skill.

As discussed in the Introduction, a previous behavioral study invoked *two* mechanisms to explain reaction time (RT) differences between reading tasks with variable attentional demands (Balota et al. 2000), suggesting that attention may interact with reading processing in more than one way. In the present report, both RT analyses (Study 1 and Study 2) and imaging analyses (Study 2) converge with the suggestion by Balota and colleagues of a 2-mechanism interaction of attentional control and reading (2000), further proposing ways in which that 2-mechanism interaction may be instantiated in the brain. A set of regions including members of the Cingulo-opercular control system and Dorsal attention system showed response properties consistent with a role in reporting a processing bias signal, reflecting pathway control, at

Mechanism #1. A set of regions including members of a left-lateralized Frontal-parietal control system showed response properties consistent with a role in response checking, at Mechanism #2. The implications of the behavioral results are discussed first, followed by a consideration of the imaging findings.

Behavioral results support the 2-mechanism model of interaction

First, large and reliable interactions of task x lexicality were observed in both Study cohorts. In both cases, the lexicality effect in regularize *reversed*, whereas Balota et al. reported a null lexicality effect (2000). Our observations therefore suggest an even more dramatic RT cost associated with lexical status during the regularize task than what Balota et al. originally noted. The 2-mechanism model depicted in Figure 2.1 suggests that attentional control interacts with reading processing at 2 loci that are relevant for word regularizing, which may explain why the task slowdown for words is so large. That is, the large reaction time cost associated with "wordness" may derive from 2 sources during the regularize task: difficulty biasing the phonological pathway at Mechanism #1, then the checking of the word's pronunciation before a response is articulated, at Mechanism #2.

Categorical treatment of frequency, using rmANOVA, revealed a null 3-way interaction of task, frequency and regularity in both cohorts, possibly reflecting the small size of any underlying interaction of those variables. Indeed, follow-up correlation analyses treating frequency as a continuous variable (which may be more appropriate, given the continuous distribution of the variable) revealed differences between the effect of frequency on regularize task performance for regular as compared to irregular words. As Balota and colleagues originally observed, the modulation of the task-induced word frequency effect reversal by regularity suggests that a mechanism other than pathway control is required to account fully for the results (Balota et al. 2000). As a whole, the behavioral results solidly thus motivated the imaging investigation of a 2-mechansism model of attentional control and reading processing. A

more detailed consideration of the behavioral results shall be provided in the context of the developmental comparison of Chapter 3.

Neuroimaging evidence supports the 2-mechanism model

Turning to Study 2, the imaging data corroborated the 2-mechanism model quite nicely. Two putative sets of control regions were identified, from among all the regions showing significant effects of (task x timecourse) or (lexicality x task x timecourse), as being likely candidates for the model proposed in Figure 2.1. Specifically, the 2 region sets identified showed distinguishable patterns of BOLD activity with regard to task, lexicality and – though it was not used to define the regions originally – regularity. The anatomical locations of the 2 clusters' respective neighbors in the dendrogram further supported the BOLD evidence for separable roles (see Supplemental figures C.3 and C.4). One cluster included 20 regions, mainly from bilateral Cingulo-opercular network and largely right-hemisphere Dorsal attention system (CO/DA+). The CO/DA+ cluster appeared to act at Mechanism #1 to bias processing along one route, depending on task demands (pathway control). The second cluster included 10 exclusively left-lateralized regions, almost all of which form part of the Frontal-parietal control network (LFP). The LFP cluster appeared to act at Mechanism #2, facilitating response verification after route processing has completed (response checking), particularly during a more difficult reading task variant.

Several observations, based on the predictions in Table 2.4, confirmed functionally separable roles for the CO/DA+ and LFP clusters and mapped those roles on the 2-mechanism model. We next discuss those key observations in detail; they are presented in summary form in Figure 2.16.

Figure 2.16 Two sets of attentional control regions map onto the 2-mechanism model
Summary schematic, based on Figure 2.1, depicting the 2 proposed mechanisms of interaction between attention and reading, along with descriptions of the region sets whose BOLD activity maps onto the distinction. A cluster of Cingulo-opercular/Dorsal attention+ regions (CO/DA+) shows activity consistent with a role in reflecting pathway control, while a cluster of left Frontal-parietal regions (LFP) shows activity consistent with a role in response checking. Compare the response properties to the predictions outlined in Table 2.4.
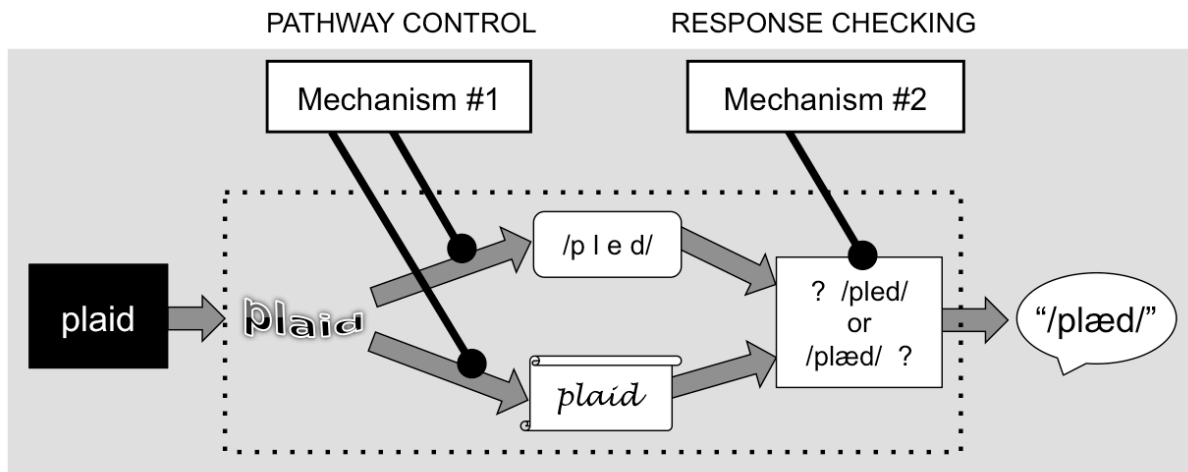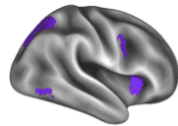
CO/DA+	LFP
• Small task effect, especially for words	• Large task effect, especially for words
• Comparably-sized lexicality effects for both tasks, in opposite directions ("opposite and equal")	• Minimal lexicality effect in read, but large (opposite) lexicality effect for regularize ("opposite and unequal")
• Less regularity-sensitive	• More regularity-sensitive
• Relatively more narrow HRF, reflecting a bias signal	• Relatively more protracted HRF, reflecting a response check
• Overall smaller BOLD responses	• Overall larger BOLD responses

CO/DA+ cluster response properties

Cingulo-opercular/ Dorsal attention+ cluster (CO/DA+) regions showed lexicality effects for the 2 tasks that were roughly "opposite and equal," with the trial-by-trial BOLD magnitudes possibly reflecting the pathway control signal relevant to a particular trial. Specifically, the CO/DA+ signals may reflect the extent to which the processing pathway most required for a task block (i.e., lexical or phonological) was successfully biased for a particular item (i.e., word or pseudoword). In this way, the signals observed in the CO/DA+ are consistent with a neural read-out of the attentional mechanism postulated by the pathway control hypothesis (as in Baluch and Derek 1991; Monsell et al. 1992; Rastle and Coltheart 1999; Zevin and Balota 2000).

The comparable magnitudes (and opposite directions) of the lexicality effects in the 2 tasks CO/DA+ regions were predicted by the fact that for each task, one of the 2 routes is heavily weighted. The regularize task clearly requires total attention to the phonological pathway output. Less obvious is that the read task, even for mixed lists (here, ~40% pseudowords), probably involves a default biasing of the lexical pathway, at least in skilled adults. In the CO/DA+ regions, the positive lexicality effect for the read task was not only equal (and opposite) to the lexicality effect for the regularize task, but it (the positive lexicality effect in read) was *larger* than the comparable effect in the LFP regions. This latter observation is consistent with the intuitive notion that for skilled adults performing normal reading, neither words nor pseudowords should require significant response checking (hence the popular notion of automaticity in reading, referenced in the Introduction).

Pathway control may be less disposable than response checking; no matter what the task variant, some signal is likely required to set the input-output parameters for task completion (Meiran 1996). Regions in bilateral aI/fO and dACC have been hypothesized to form the "core" of the task-set system (Dosenbach et al. 2006), although their precise functional ascription remains elusive (Posner and Petersen 1990; Dehaene et al. 1994; MacDonald et al. 2000; Bush

et al. 2000; Botvinick et al. 2004; Brown and Braver 2005). Because the most consistent control-related characterization of the core Cingulo-opercular regions pertains to their involvement in sustained, task set-maintenance (and this event-related study design did not measure the task-related sustained signal; but see Chapter 4 for a future direction), the exact specification of the role of the *trial-by-trial* BOLD signals observed here entails some amount of speculation. Nonetheless, the importance of the core Cingulo-opercular regions, bilateral aI/fO and dACC, in shaping task performance is underscored by the observation that they tend to be among the most commonly activated regions across all kinds of neuroimaging studies (Nelson et al. 2010). It is also interesting that neuroimaging studies of the Stroop effect have noted strong and reliable activations in the dACC in relation to word-color interference (e.g., Pardo et al. 1990; Bench et al. 1993; Carter et al. 1995; Peterson et al. 1999; see also Bush et al. 1998 for a related, though not identical, phenomenon).

Signals reporting pathway control ought not to reflect appreciably the effect of word regularity, since both regular and irregular words are thought to generate activations along both pathways obligatorily (Coltheart et al. 2001). Indeed, CO/DA+ regions showed smaller effects of regularity than LFP regions, considering both the read task and the regularize task. Examined another way, the CO/DA+ regions also showed smaller effects of task when each of regular and irregular words was considered separately. Note, however, that the CO/DA+ regions *did* show some sensitivity to regularity; in each task separately, irregular words produced significantly larger activations than regular words. Given that regions in the Cingulo-opercular network simultaneously carry several types of processing signals (Neta et al. *under revision*), perhaps the relatively small effects of regularity observed in the CO/DA+ regions reflect an orthogonal response property, such as time-on-task. The important point is that the effects of regularity are much more pronounced in the LFP regions than the CO/DA+ regions.

In addition to Cingulo-opercular regions, the CO/DA+ cluster includes several regions of the Dorsal attention system, a control system implicated in the allocation of top-down attention control to specific portions of visual space (e.g., Corbetta and Shulman 2002). Resting-state functional connectivity (rs-fcMRI) work from our own laboratory has specifically suggested that Dorsal attention system regions are more closely functionally coupled to a region in left ventral occipital temporal cortex than are classic reading processing regions (Vogel et al. 2012), underscoring the relatedness of the type of processing performed by Dorsal attention system regions and the type of processing required for orthographic analysis. Our finding of a close functional relationship between Cingulo-opercular regions and Dorsal attention regions as part of a single cluster is consistent with other work suggesting that the Dorsal attention system may flexibly modulate its coupling with other networks, depending on task demands (Spreng et al. 2010). Possibly the regions in the CO/DA+ cluster work together to read out the pathway control signals during the read/regularize tasks through a combination of each network's unique processing capabilities. It would be interesting to see whether for *other* reading task variants – perhaps a task in which serial, left-to-right attention to phonological decoding were not a critical component, as it is in regularize – would result in the reporting of pathway control signals by Cingulo-opercular regions working less closely with or independently from Dorsal attention system regions.

LFP cluster response properties

Regarding left Frontal-parietal cluster (LFP) regions, it was predicted that task, both as a main effect and as interacting variable, would exert a larger influence at Mechanism #2 regions than at Mechanism #1 regions. Indeed, there was a larger task effect for the LFP regions than the CO/DA+ regions, for words specifically. For pseudowords, the 2 regions sets showed statistically indistinguishable sensitivity to task. That is, the task x timecourse x cluster interaction was not significant for pseudowords, although a significant effect of task for

pseudowords was separately obtained in the CO/DA+ regions, but not the LFP regions. The fact that the clusters were distinguishable by their task effects for words, but not pseudowords, was neither explicitly predicted nor surprising. First, the reaction time task effect was much larger for words than for pseudowords, and fMRI effects are generally weaker when the corresponding behavioral effects are weaker. Furthermore, although response checking is extremely important for the regularize task, that which makes it especially tricky – familiarity with both a word-form and the process of response checking for that word-form – does not factor in for pseudowords. However, as shown in Figures 2.11 and 2.12, the timecourse of signal change for pseudowords in the regularize task was still large and positive in the LFP cluster (and in fact larger in terms of % signal change than the corresponding response in the CO/DA+ cluster), suggesting that pseudowords may indeed undergo response checking (perhaps equally so for the read and regularize tasks).

Likewise consistent with a role at Mechanism #2, LFP regions showed unequal regularity effects in the two tasks – in both cases irregular > regular words – with the effect being larger in the regularize as compared to the read task. The larger impact of regularity in the regularize task may reflect the discrepancy between the relative ease of response checking for normal reading (even for irregular words, as previously discussed) vs. the increased need for response checking for the regularize task. Indeed, regions in dorsolateral prefrontal cortex (dlPFC) and posterior parietal cortex (PPC) – possibly overlapping with some of the members of the LFP cluster in the present results – have been hypothesized in other contexts to mediate dynamic stimulus-response mappings (Huettel et al. 2005), perhaps encompassing the kind of computation required to generate /pInt/ (like "lint") in response to seeing "pint." Relatedly, individual FP regions in both inferior prefrontal cortex (IPC) and dorsolateral prefrontal cortex (dlPFC) have been implicated in various attention-related processes (Norman and Shallice 1986) including response retrieval (Wagner et al. 2001), response checking (Thompson-Schill et

94

al. 1997), and the maintenance of stimulus-response mappings (MacDonald et al. 2000).

Frontal-parietal regions in posterior parietal cortex (PPC) have been noted for their importance

in facilitating decision-making in support of task demands (Platt and Glimcher 1999; Rushworth

et al. 2001; Bunge et al. 2002). All of these functional ascriptions are consistent with a role for

left Frontal-parietal regions at Mechanism #2 in selecting and verifying the appropriate response

from between the 2 route computations.

Parallel to the argument for unequal regularity effects, LFP regions showed lexicality

effects that were "opposite and unequal," in contrast to those observed in CO/DA+ regions

("opposite and equal"). In other words, the effect of lexicality was larger in LFP regions for the

task requiring the non-practiced response checking (regularize). Also, the negative lexicality

effect in the LFP regions for regularize was larger than the corresponding effect in the CO/DA+

regions.

Together with the large task effect for words and null task effect for pseudowords, the

relatively larger negative lexicality effect for regularize observed in the LFP region set suggests

that the LFP regions may actually perform *more* of the processing associated with the regularize

task, compared to the CO/DA+ regions. In other words, a (speculative) argument could be made

that the 3 factors just described – large task effect for words, null task effect for pseudowords

and large negative lexicality effect for regularize – index the relative importance of a region's

contribution to the attentional control required for the regularize task. One additional hint of the

primacy of the LFP regions in directing attentional control during reading is that the effects in the

LFP cluster seemed slightly more robust to particular analysis choices, including reaction time

regression, than are the effects in the CO/DA+ cluster (see Supplemental materials A.3a and

A.3c). Large and robust task-sensitive responses in the LFP regions may reflect the extra-lexical,

domain general nature of the processing in those regions, as characterized, for example, in one

dlPFC region (Tal -50, +23, +27; MNI -53, +26, +30) in a previous study (Fiebach et al. 2007).

95

Magnitude, timing and HRF shape differences between the clusters

Bolstering the hypothesis that the processing performed by the LFP regions is particularly critical to the attentional control of reading is the set of general observations about the features of the clusters' BOLD responses. Specifically, the CO/DA+ and LFP clusters differed in terms of overall BOLD magnitude, time-to-peak and width of the HRF. Compared to the CO/DA+ cluster, the LFP cluster (both in aggregate and in terms of the individual regions) showed responses that were slower to rise, higher to peak and slower to fall. Because of the sluggishness of the BOLD response and the timescale over which the data is collected, presumptive timing differences observed using standard event-related fMRI should be considered speculative. However, the relatively large, relatively wide HRFs of the LFP regions are internally consistent with the argument articulated above, that the LFP regions perform something of a *primary* role in the trial-by-trial attentional control required for the read and regularize tasks. It is also intriguing to consider the possibility that the wide HRFs seen in the LFP regions reflect the linear summation of more than one BOLD response within a single trial (Boynton et al. 1996), i.e., an even more complex role for the LFP regions in coordinating the control of reading than that which is suggested by the model in the summary Figure 2.16.

A final general observation about the 2 clusters is that the LFP cluster includes only regions in the left hemisphere, whereas the CO/DA+ cluster includes regions that span the right hemisphere, left hemisphere and midline. We did not predict that the task-related control signals would segregate into 2 clusters with such a dramatic difference in hemispheric lateralization, and it is unclear what the left-hemisphere exclusivity of the LFP cluster means in the context of read and regularize. Of course, it has long been posited that language relies primarily on left hemisphere structures, but indeed this and most other imaging studies in the recent literature show that brain activity related to reading and language tasks is very much dependent on right-hemisphere regions as well (e.g., Jobard et al. 2003; Taylor et al. 2012; Price 2012). Future

studies could interrogate possible task differences in the extent to which Frontal-parietal and other control region networks are engaged either bilaterally or relatively unilaterally, and what such differences may mean in terms of information processing.

Both sets' neighbors in the dendrogram

The profiles of the clustering neighbors of the CO/DA+ and LFP region sets strengthened their respective proposed roles at Mechanism #1 and #2. The CO/DA+ regions showed neighborhood relationships consistent with a more intimate interaction with the earlier stages of visual word recognition, i.e., a role at Mechanism #1 (pathway control). For example, they appeared close to bilateral middle occipital (visual) regions and some additional Dorsal attention regions (medial and superior frontal) in the clustering dendrogram. The most proximal neighbors to the LFP regions were left and right Inferior frontal gyrus *pars opercularis* IFG*po*. A region in left IFG*po* has been consistently implicated in preparing the articulatory response during reading (e.g., Jobard et al. 2003; Taylor et al. 2012). Regions in bilateral cerebellum as well as bilateral mouth somatomotor cortex were also relatively close to the LFP regions, as were regions in bilateral superior temporal gyrus. This clustering similarity suggests a tighter functional link between LFP regions (compared to the CO/DA+ cluster) and regions important for speech articulation (Price 2012), corroborating a role for the LFP cluster at Mechanism #2 (response checking) in the proposed model.

Contextualizing these results in the literature

Despite very different analysis approaches, there is some convergence between the results of the present study and the findings of the only extant neuroimaging analysis of the regularize task of which we are aware (Gold et al. 2005). Gold and colleagues asked whether the often-cited anterior-semantic/posterior-phonological distinction along the left inferior prefrontal cortex (LIPC) (Demonet et al. 1992; Buckner et al. 1995; Fiez 1997; Poldrack et al. 1999; McDermott et al. 2003) might be better captured as the presence vs. the absence of

controlled processing, since many previous comparisons between semantic and phonological tasks were confounded by greater demands for controlled processing in the semantic tasks. Using 2 tasks that both require controlled processing (regularize-phonological, Verb generation-semantic), those authors indeed observed BOLD activity effects in *both* anterior and posterior portions of LIPC during *both* tasks, which they interpreted as evidence for a more general role for LIPC in task-driven attentional control of verbal information (Petrides et al. 1995). Instead of observing domain specificity in LIPC, they noted 2 other regions, out of LIPC, that showed preferential activity for one of the 2 tasks. A left frontal region (BA 6, Tal -55, -1, +28; MNI -58, 0, +31) showed BOLD effects that were larger for the regularize task than the Verb generation task, while a left temporal region (BA21, Tal -51, -55, +2; MNI -54, -58, +3) showed the opposite pattern. The results of the present study are broadly consistent with the Gold et al. conclusion that "reading"-related activity, in left frontal regions in particular, is modulated in important ways not only by the kind of processing required (e.g., sound-based vs. meaning-based vs. orthography-based) but by the extent to which that processing requires attentional control.

Other neuroimaging investigations of the attentional control of reading have generated intriguing results, but with very different angles of approach from the read/regularize comparison employed here. For example, one study isolated "top-down" effects from "bottom-up" effects in individual letter recognition by training subjects to "detect" letters from noisy arrays, where no such letters existed (Liu et al. 2010). More often, reports on "top-down" effects in reading consider attention effects at the opposite end of the spectrum, i.e., resulting not from strategic and conscious allocation of attention but instead from experimental manipulations designed to reveal aspects of the intersection of attention and reading that occur outside of conscious awareness. For example, important insights regarding the bidirectional flow of information between "higher" and "lower" –level reading processing regions have derived from close examination of intrinsic stimulus properties (Devlin et al. 2005), priming effects (Nakamura et al.

2007; Kherif et al. 2011) and clever timing manipulations aimed at separating perceptual from linguistic processes (Davis et al. 2011).

Some studies have employed paradigms requiring a conscious reallocation of attention more similar to the task comparison introduced by Balota et al. (2000) and replicated here. For example, some electrophysiological data suggests that different task instructions can modulate word identification as shown, for example, by an enhanced N200 response during an orthographic, as compared to semantic and phonological, task (Ruz and Nobre 2008), and by distinguishable patterns of preparatory brain activity in anticipation of different task requirements (Ruz and Nobre 2008). The greater spatial resolution of PET and fMRI have also rendered informative observations, but many of these analyses have focused on reading-related and primary sensory regions (e.g., Roskies et al. 2001; McDermott et al. 2003; Dietz et al. 2005; Yoncheva et al. 2010; Twomey et al. 2011; Perrone-Bertolotti et al. 2012). The tendency for previous investigations to emphasize presumptive effects of controlled attention in processing regions, rather than focusing on the control signals themselves, may reflect the greater transparency of the predictions of the dual route and other models with regard to processing as compared to attentional control (for a recent meta-analysis, see Taylor et al. 2012). This is a speculation that echoes the discussion of the automaticity dogma in Chapter 1.

The present results add to the body of literature on the attentional control of reading by focusing explicitly on the contributions of signals from previously described control-related regions, using targeted behavioral results to test neuroimaging hypotheses. Specifically, we examined whether and how task instructions requiring a novel stimulus-response mapping for an otherwise automatically generated output (i.e., the *conscious* reallocation of selective attention) modulates both behavior and brain activity. Our data suggest that skilled adult readers *can* strategically direct attention to one component of reading, phonological processing, but with much effort (in terms of slow reaction times). Critically, both the behavioral and the imaging data

are best accounted for by a model in which 2 distinct attentional control processes interact with reading processes, both to report the bias signal associated with pathway control (CO/DA+ regions) and to ensure that the correct response is selected (LFP regions).

The observations of the current study may provide a framework for future investigations of dyslexia, since dyslexia is disproportionately comorbid with ADHD (Willcutt et al. 2010; Willcutt and Pennington 2000; Germano et al. 2010), and thus far, few studies have explicitly examined whether brain activity in attentional control-related regions may help explain the pathophysiology of this most prevalent learning disorder (but see Eden et al. 1996; Eden et al. 1996; Demb et al. 1998; Bosse et al. 2007; Vidyasagar and Pammer 2010; Peyrin et al. 2012; Koyama et al. 2013 for some related findings). As will be discussed in Chapter 4, future studies could also include extending the analyses described here to a population of non-impaired, school-age, beginning readers, to investigate how reading skill modulates both task performance and control region recruitment. For example, it is likely that children will show smaller reaction time task x lexicality interactions than adults, and possibly also smaller effects of regularity on the task-induced word frequency effect change. It would also be interesting to investigate in children whether regions in the Cingulo-opercular and Dorsal attention systems would show functional similarity to each other, and functional dissimilarity from left Frontal-parietal regions, as was observed here in adults. If a CO/DA+ vs. LFP or some other distinction were observed in children, one follow-up question would be whether evidence could be found for a differential timecourse of development for these two sub-systems and their point of interaction with reading processing regions.

CHAPTER 3: TWO MECHANISMS OF ATTENTIONAL CONTROL OF READING-RELATED TASK PERFORMANCE SHOW DEVELOPMENTAL EFFECTS

S. Katie Ihnen, Rebecca Treiman, David A. Balota, Steven E. Petersen and Bradley L. Schlaggar

Abstract

How the resource-demanding task of learning to decode single words eventually gives rise to the effortless extraction of meaning from connected text remains a central and unanswered question in the study of the development of skilled reading. This study investigated the role of attentional control in the reading of single words across 2 age groups, building on Chapter 2 examining young adults. Children (8-10 years old) and young adults (22-28 years old) performed two speeded naming tasks, one requiring standard pronunciation of single words ("read") and the other requiring selective attention to the phonetic pronunciation (e.g., pronounce "pint" such that it rhymes with "lint;" "regularize") (adapted from Balota et al. 2000). A reliable crossover interaction was observed in which adults were (as expected) faster than children in the read task, whereas children were (somewhat unpredictably) faster than adults in the regularize task. Age group effects were also observed with regard to how the read and regularize tasks were differentially affected by stimulus characteristics. *Post hoc* correlation analyses suggested that the word frequency effect was modulated by regularity, task and age group. Additionally, a 3-way interaction of task, lexicality and age group revealed that adults, compared to children, were particularly slow to regularize words. Critically, the task effect for pseudowords, like the task effect for words, was significantly greater in the adults than in the children. Together, the findings corroborate an earlier suggestion (Balota et al. 2000) of at 2 stages, or mechanisms, of interaction between attentional control and reading processing, specifically, pathway control and goal-directed response selection. The results further suggest

that the development of skilled reading entails changes in both of the proposed attentional

mechanisms, a tentative hypothesis that could be tested by future imaging analyses. These

results underscore the importance of explicitly accounting for attentional effects when

interpreting performance on reading-related tasks, in both beginning and experienced readers.

Introduction

Much of the most recent research on reading has been influenced by a set of

computational models that reflect a range of architectural heritages (e.g., Seidenberg and

McClelland 1989; Coltheart et al. 1993; Norris 1994; Zorzi et al. 1998; Perry et al. 2007). These

computational models of reading have been repeatedly and impressively refined in the service

of accounting for experimentally-observed phenomena related to visual word processing,

simulating especially well the impact of lexical characteristics on speeded naming and lexical

decision task response latencies (e.g., Coltheart and Rastle 1994; Coltheart et al. 2001; Harm

and Seidenberg 1999; Harm and Seidenberg 2004; Grainger and Ziegler 2011; Norris and

Kinoshita 2012; Pritchard et al. 2012; Perry et al. 2013). One particularly influential class of

models is referred to as "dual route," positing 2 processing avenues of print-to-sound translation

(e.g., Coltheart et al. 1993; Coltheart et al. 2001).

According to dual route theory, skilled reading involves parallel activation along both the

phonological (serial grapheme-to-phoneme conversion) (Forster and Davis 1991) and lexical

(whole-word look-up) routes, with the lexical route generally completing its processing more

quickly than the phonological route, at least for skilled readers presented with familiar,

frequently occurring words (Frederiksen and Kroll 1976; Paap and Noel 1991; Visser and

Besner 2001). While the outputs from the two routes agree for regular words (e.g. "cat"), they

produce competing outputs (and thus slower response latencies) for words with irregular

spelling-to-sound correspondences such as "plaid" (for which the phonological route generates

the pronunciation /pled/). Whether or not the 2 route computations converge is determined once

the "full reading process" has occurred, i.e., after route computation (Taylor et al. 2012).

Pseudowords (pronounceable though unfamiliar letter strings) can only be *fully* decoded via the

phonological route. Although explications of dual route frameworks tend to focus on the

processing that occurs along the 2 routes, a sometimes-overlooked final step of "phoneme

recognition" is necessary to select the response from between the 2 possibilities generated by

the routes (Pritchard et al. 2012), e.g., selecting /plæd/ rather than /pled/.

Generally, the dual route and other computational models have pursued the objective of

describing the component processes of reading (e.g., phonological, orthographic and semantic

analysis) in the absence of an explicit consideration of the role of attentional control in shaping

these processes and their necessary integration (see McCann et al. 2000; Reynolds and Besner

2006 for further discussion). That is, many of the studies describing computational models of

reading have discussed results in a way that implies relative independence of reading from

attentional control mechanisms. The empirical research on reading has generally mirrored this

pointed focus, conceptualizing reading as an encapsulated set of processes immune from the

purview of attentional control (exceptions do exist, as will be discussed).

One plausible explanation for the apparent paucity of interest in the intersection of

attentional control and reading processing may relate to the assumption that reading, at least in

skilled adults, is an automatic process. The supposition of automaticity in reading is grounded

both in important theoretical considerations (e.g., LaBerge and Samuels 1974; Stanovich 1990)

and in compelling empirical evidence for the simultaneous and obligatory activation of multiple

word representations, regardless of task (ir)relevance (as shown prominently by Stroop 1935;

see MacLeod 1991 for a review of Stroop studies; for other paradigms, see Neely 1977; Luck et

al. 1996; Perfetti 1992; Booth et al. 1999; Booth et al. 2008). But automaticity is notoriously

difficult to define e.g. (e.g., Stanovich 1990; Logan 1997) – by one account requiring

consideration of 2,047 permutations of various criteria (Reynolds and Besner 2006).

Furthermore, whether or not skilled reading in adults can properly be considered automatic

(which is perhaps more appropriately considered to be a matter of degree, rather than kind (as

suggested by Logan 1997)), important orthogonal questions remain. For example, at what point

developmentally and under what circumstances in adulthood does reading processing depart

from automaticity? Which components of reading are especially susceptible to attentional

control, and with what developmental trajectory? To what extent does a reader's attentional

demand vary across a given text? How do brain regions important for attentional control interact

with reading "processing" regions?

      A relatively small number of studies in healthy young adults have addressed attentional

control-related effects in speeded pronunciation tasks. For example, list composition

manipulations designed to cause subjects to favor processing along one route (say,

phonological) at the expense of the other have demonstrated that reaction times to particular

item types can increase or decrease depending on which route has been more heavily utilized

by recently presented items (e.g., Baluch and Derek 1991; Monsell et al. 1992; Rastle and

Coltheart 1999; Reynolds et al. 2012). Those studies and others using related paradigms (e.g.,

priming, as in Zevin and Balota 2000) have generated fruitful debates regarding the interplay of

lexical and phonological processing, inspiring the elaboration of heuristics including, prominently,

the pathway control hypothesis (e.g., Monsell et al. 1992; Zevin and Balota 2000), and the

response time homogenization hypothesis (e.g., Lupker et al. 1997; Kinoshita and Lupker 2003).

The pathway control hypothesis – the idea that skilled readers can exert attentional control over

reading by differentially emphasizing the relative contribution of either the phonological or lexical

route – is particularly relevant to the current study, as discussed with regard to the model

proposed in Chapter 2.

List composition and priming studies primarily invoke implicit reallocations of attention, therefore relying upon at least some small set of assumptions regarding the efficacy of the task and/or stimulus manipulations. Comparatively few studies have utilized *explicit* attentional manipulations during reading to interrogate more directly the relationship between goal-directed attention and reading processing. One example of a direct manipulation of attention in reading involved young adults performing two tasks, a simple speeded naming task and a "sound it out" task (e.g. pronounce "pint" (/pajnt/) to sound like "lint" (/lInt/), thus directing attention only to phonological processing) (Balota et al. 2000). Considering a dual route framework, Balota, Law and Zevin reasoned that if skilled readers can consciously direct attentional control *completely* to *frequency-independent* grapheme-to-phoneme conversion to "sound out" (regularize) words, then the typically observed main effects of lexicality and word frequency (suggestive of lexical route activation) should be reduced in the regularize task as compared to the read task.

The Balota et al. study furnished several key observations. Not surprisingly, subjects performed the regularize task more slowly than the read task, by several hundred milliseconds. Specifically, participants were especially slow to regularize words, losing completely the reaction time advantage they showed for words compared to pseudowords in the read task. The reliable interaction of task by lexicality suggested that readers failed to inhibit completely the lexical pathway for words, despite task instructions to focus on the sounds. Additionally, the typically observed word frequency effect (WFE; high frequency words are pronounced more quickly than low frequency words) *reversed* in the regularize task for regular, but not irregular words. This final observation yielded a significant interaction of task, frequency and regularity, and was interpreted as evidence of more than one mechanism of attentional control for reading. Specifically, Balota and colleagues proposed that some sort of response checking step is implemented prior to articulation, during the regularize task. Those authors proposed that a particularly prolonged response check may be incurred on some portion of the high frequency

105

regular word trials, due to the reader second-guessing the (normal) pronunciation, subconsciously thinking the pronunciation should sound "funny."

The findings of the Balota et al. study regarding the attentional control of reading in healthy young adults raise provocative questions about the role of goal-directed attention in early readers. With the present study, we sought to examine some of those developmental questions. Our first primary question was the following: do school-age children show a larger, smaller or comparable RT cost compared to young adults, when instructed to direct attention entirely to phonological processing? Essential priors make competing predictions with regard to the overall task x age group interaction. On the one hand, (1) children are slower than adults to perform almost all cognitive tasks (e.g., Kail 1991), and (2) in particular, children are typically less able than are adults to exert selective attention (e.g., Rueda et al. 2004; Luna et al. 2004; Zelazo et al. 2004; Davidson et al. 2006), as required for pathway control. These two general developmental premises predict slower performance by the children than the adults on both the read and regularize tasks, at least in terms of absolute speed.

On the other hand, children are likely to have weaker lexical representations than adults (e.g., Zoccolotti et al. 2005), directly reflecting their relatively less extensive exposure to written word forms (Stanovich 1993; Share and Stanovich 1995). Weaker lexical representations may be implemented in a dual route framework as either higher thresholds of activation (Andrews 1989) or lower baseline activation levels (Morton 1969) for individual lexical entries. This premise, related to experience rather than development, predicts that children should perform the regularize task faster than adults, due to a reduced contribution from the interference of lexical representations, at least relative to their respective performance on the read task.

Unfortunately, both predictions (children will be slower/faster than adults to regularize) are not so logically straightforward as was just implied. More realistically, the cognitive development-related and experience-related factors that distinguish children from adults will

106

exert opposing effects on pathway control, with the children, compared to the adults, showing *both* reduced ability to exert inhibitory control *and* weaker lexical representations requiring inhibition in the first place. Because of the push/pull between these factors, the outcome of the task x age group interaction will depend upon whether immaturity of selective attention in the children (e.g., Rueda et al. 2004) or strong, consolidated lexical representations in the adults (e.g., Coltheart et al. 2001) exerts a dominant influence.

Complicating the issue further is a consideration of 2 additional potential age group differences that may interact with lexical pathway control to drive a task x age group interaction. First, the ability to compute sublexical spelling-to-sound conversions will likely influence overall regularize task performance. Arguably, 8-10 year old children rely heavily on phonological processing (Backman et al. 1984; Aaron et al. 1999; Pugh et al. 2001; Greenberg et al. 2002), possibly moreso than adults (Doctor and Coltheart 1980; Greenberg et al. 1997), especially given that children in this age range decode an average of 4,000-12,000 new words per year (Scholastic: www.**scholastic**.com/readeveryday/pdfs/reading_**facts**.pdf*).* Then again, the highly educated young adults in our study sample have had many more years of cumulative experience with spelling-to-sound mappings than have the children, and some evidence suggests automatic engagement of phonological processing in adults, regardless of task demands (e.g., Perfetti 1992; Booth et al. 1999; Booth et al. 2008; Cone et al. 2008). Yet another potentially contributing factor is that the quality and quantity of response checking is likely to matter significantly in overall task performance. Indeed, it may be that response checking plays a larger role in the attentional control of regularize task performance than pathway control, at least in young adults, as suggested by the observations in the imaging portion of Chapter 2 showing greater overall activity and greater modulation by stimulus features in the left Frontal-parietal (LFP) as compared to the Cingulo-opercular/ Dorsal attention + (CO/DA+) cluster.

While it is difficult to tease out the distinct contribution of age group differences in grapheme-to-phoneme conversion ability in the current design, we may be able to identify age group differences in the response check step. Although the current study does not include neuroimaging and therefore does not allow for a comparison of the relative contributions of LFP and CO/DA+ activity in the 2 age groups (see Chapter 4 for a future experiment), the present data can be leveraged to examine reaction time differences due to stimulus features, consistent with the logic of Balota and colleagues (2000). Thus, our second primary question was as follows: do children and adults show differences in the way that stimulus characteristics interact with task, in a way that may distinguish developmental differences in pathway control from developmental differences in response selection?

As was the case with the pathway control predictions, a prediction for an age group difference in response checking should be informed by priors. One relevant experience-related factor is that the children should have relatively sparser meta-linguistic knowledge (Nagy and Anderson 1995)[7]. Secondly, and related to cognitive development writ large, 8-10 year old children are yet immature with regard to the kind of attentional control that may be needed to implement a response check (Norman and Shallice 1986), as evidenced, for example, by inferior ability to monitor their responses (Kopp 1982; Davies et al. 2004). Thus, one prediction regarding our second question is that response checking during the regularize task should be less contributory to the performance of the children than the performance of the adults. If children engage response checking less frequently and/or more briefly than do adults, we should see weaker interactions, in the younger group compared to the older group, of stimulus characteristics with task. For example, the children may show a smaller regularize task-induced reversal of the word frequency effect than the adults, particularly for regular words. This

---

[7] Meta-linguistic knowledge captures that level of familiarity with words, spellings, sounds and meanings that likely produces or prolongs a decision heuristic such as a response check.

prediction follows from the original Balota et al. interpretation of the task x frequency x regularity interaction in adults as evidence for the response checking step in the first place.

Another approach to interrogating age group differences in response checking is to examine the interaction of task, lexicality and age group. A 3-way interaction of those variables that is characterized by an age group difference in the task effect for words but not for pseudowords suggests an age group difference in pathway control, since the words, but not the pseudowords should require inhibition of the lexical pathway. On the other hand, a 3-way interaction that is characterized by age group differences in *both* the task effect for pseudowords and the task effect for words points to age group differences in both pathway control and response checking, since pseudowords lack lexical representations, and their pronunciations, if anything, should be facilitated by selective attention to phonological processing during the regularize task.

Building off of Chapter 2, the data presented in this chapter describes a behavioral comparison of the regularize vs. the read task in a group of 8-10 year old children and a group of young adults. We asked 2 primary questions in this chapter. First, do school-age children show a larger, smaller or comparable reaction time cost compared to young adults, when instructed to direct selective attention to phonological processing? Second, do children and adults show differences in the way that stimulus characteristics interact with task, in a way that distinguishes developmental differences in pathway control from developmental differences in response selection? In asking these 2 primary questions, we pursued 2 objectives: (1) to re-asses, in 2 cohorts of different ages and reading levels, evidence for the 2-mechanism model of attentional control in reading, as described in Chapter 2; and (2) to characterize any age group differences illuminated by the first objective, as a way of elucidating the role of attentional control in reading development. As presaged, the results discussed in this chapter will

eventually be followed by a corollary neuroimaging investigation of the age group differences that are revealed.

Method

Participants

Participants were recruited from the Washington University community and elementary schools in the area. Any potential participant with neurologic and/or psychiatric diagnoses, any history of dyslexia or reading difficulty, or current use of psychotropic medications was not included in the study. Enrolled participants included 32 adult (ages 22-28 years; 15 male) and 35 child (ages 8-10 years; grades second through fifth; 17 male) right-handed, native and monolingual English speakers. Data from 4 children were excluded, 3 because they failed to perform both tasks at the accuracy criterion of >65%, and 1 who failed to meet the criterion for accuracy on the post-test, see below. Thus the analyzed data derives from 32 adults (15 male) and 31 children (16 male).

Neuropsychological tests were administered to each participant. Two subtests of the Wechsler Abbreviated Scale of Intelligence (Vocabulary and Matrix Reasoning) assessed general intelligence (Wechsler 1999). Three subtests of the Woodcock–Johnson III (Letter–Word ID, Reading Fluency, and Word Attack) estimated absolute and relative reading level (Woodcock and Johnson 2002). A Stroop color naming task assessed attentional control (Spieler et al. 1996). Supplemental tables D.1 and D.2 summarize the groups' neuropsychological testing results. Two-tailed, independent samples t-tests between the adults and the children were conducted for each cognitive test, and the resulting test statistics are also reported in the tables.

In addition to establishing that the groups were comparably skilled for grade-adjusted reading level (see Supplemental material D.3 for more information), neuropsychological testing

established that all participants were of at least average intelligence. While there were reliable

age group differences in some testing results, many individual scores, even of those measures

showing a reliable group difference, did not map onto task performance. For example, the

groups were different on several measures of Stroop performance: children were slower (for

each of the 3 conditions) and less accurate (for incongruent and neutral trials) than the adults. A

normalized index of Stroop interference was calculated for each subject (incongruent median

zRT - congruent median zRT). There was a significant age group difference in this measure,

with children showing greater interference than adults (see Supplemental table D.2). However,

there was not a reliable correlation between Stroop interference and any index of read and

regularize task performance. See Supplemental material D.2 for a more detailed discussion of

individual and age group Stroop results.

Some neuropsychological scores did reliably correlate with individual task performance

measures of interest (task measures described later). Although neuropsychological correlation

analyses were not the focus of this chapter, Supplemental material D.4 describes several

noteworthy correlations in the test scores, the implications of which are addressed in Chapter 4

in the general Conclusions.

Procedure

Subjects were seated in front of a 16 inch Sony CRT computer monitor, with a standing

microphone placed 2-3 cm away from the mouth. Stimuli were controlled using Psyscope X B53

on a Power Macintosh G3 (Cohen et al. 1993). Vocal responses were transmitted to a Sony

MP3 Digital Voice Recorder (ICD-UX70 1GB USB).

For the regularize task, participants were asked to "sound out" items as though reading

them for the first time, using whatever strategies they normally use when encountering new

words. Thirteen demonstration trials and 73 practice trials of the regularize task were then

presented (see Supplemental materials B.5a and B.5b for practice items, including example

pronunciations). The practice trials were self-paced with feedback; correct answers were given by the experimenter as needed. No practice items were repeated in the tasks.

The experimental session began with 3 consecutive blocks of the regularize task, followed by 3 consecutive blocks of the read task ("read normally."). Each block included 106 items and lasted 6 min, 10 s. Pilot testing suggested that the fixed regularize/read ordering was most effective for establishing and maintaining task mode.

Stimuli were displayed on a black background in lowercase white letters at font size 56, subtending 2.1 to 5.0 degrees of visual angle horizontally and 0.7 to 1.1 degrees vertically. Each stimulus appeared alone, and in the middle of the screen, for 2500 ms. Participants were free to respond as soon as the stimulus appeared and up until the next stimulus appeared. Immediately following the disappearance of each stimulus, a white fixation crosshair appeared in the middle of the screen for 1 s. Participants were asked to maintain visual fixation on the screen for the duration of the block.

After completing the regularize and the read tasks, participants performed a 2-step lexical decision/semantic knowledge post-test to ensure familiarity with the words. The 636 stimuli were divided into 6 lists (not the same lists used in the tasks), and list order was counterbalanced across participants. First, an item appeared on the computer screen along with the choices "WORD" and "NOT A WORD" (the same projection and display equipment was used for the post-test as for the tasks). Subjects pressed the right button of a Psyscope-compatible button box (Cohen et al. 1993) for one option and the left button of a button box for the alternative (button assignment was also counterbalanced across subjects). For those words that were correctly identified as words, three short definitions subsequently appeared on the screen. The participant pressed a button corresponding to the best choice among the options. For the pseudowords and for the words that were incorrectly identified as non-words, no definition choices were given. The post-test was untimed.

Adults made few errors on the post-test (>99% accuracy); children achieved, on average, 92% accuracy. One child failed to meet the 65% criterion and was excluded from the task data.

Stimuli

The stimuli included 252 pronounceable pseudowords and 384 words. Supplemental material B.1 includes the complete list of test stimuli, and Appendix B provides additional stimulus information. The words included 96 of each of 4 combinations of regularity (regular or irregular) and frequency (high or low). Items were 3-8 letters long, with an average length of 5.6 letters and 1.5 syllables. Table 3.1 shows the mean values of several lexical properties for each of the 5 stimulus classes.

Word frequency ratings were based on the Zeno et al. standard frequency index (SFI) (1995). Stimuli spanned an SFI continuum from 40.0 to 75.0 (unit = log-transformation of a weighted frequency per million; see Supplemental figure B.8 for a depiction of the distribution across the frequency range). A median split designated subsets of equal numbers of regular and irregular words as high frequency (M SFI = 60.0) and low frequency (M SFI = 47.0). High-frequency words had significantly higher SFIs than low-frequency words (t(349) = 29.128; p < .001. Regular and irregular words *within* each frequency category were not statistically different from one another in terms of SFI (low frequency: t(190) = 0.907; p = .365; high frequency: t(184) = -1.953; p = .052).

A word was determined to be regular if its rime (e.g. –at in cat) followed common spelling to sound correspondences. For example, "gave" (/gev/) is regular whereas "have" (/hæv/) is irregular, because the "e" at the end of "have" should generate the pronunciation /hev/. Pilot testing confirmed that words designated as irregular typically resulted in different pronunciations for the 2 tasks, while regular words did not. Pseudowords were created using onsets (e.g. /k/ in "cat") and rimes (e.g. /æt/ in "cat") that tend to be have unambiguous pronunciations).

Table 3.1[8] **Mean values of lexical properties across 5 stimulus types (short and long)**

Stimulus type[a]	Number of letters (SD)	Number of syllables (SD)	Number of orthographic neighbors[b] (SD)	Number of phonological neighbors[b] (SD)	Standard frequency index[c] (SD)
High frequency regular words	5.51 (1.45)	1.50 (0.50)	4.52 (5.12)	8.64 (9.93)	59.34 (4.51)
Low frequency regular words	5.52 (1.45)	1.50 (0.50)	4.14 (4.98)	8.39 (10.11)	47.22 (3.83)
High frequency irregular words	5.56 (1.41)	1.50 (0.56)	3.31 (4.10)	8.99 (9.56)	60.75 (5.42)
Low frequency irregular words	5.52 (1.44)	1.53 (0.56)	3.08 (4.44)	8.36 (10.48)	46.74 (3.48)
Pseudowords	5.52 (1.42)	1.50 (0.51)	2.98 (4.03)	---	---

Note. [a]For all word stimulus types, n = 96; for pseudowords, n = 252. [b]Orthographic neighbors and phonological neighbors calculated using the English Lexicon Project database (http://elexicon.wustl.edu; Balota et al. 2007). [c]Standard frequency index (SFI) is from Zeno et al. (Zeno et al. 1995).

As shown in Table 3.1, the stimuli were well matched for number of letters, number of orthographic neighbors and number of phonological neighbors. Although not indicated in the table, stimuli were also selected such that acoustic properties of the onset phoneme did not systematically vary by type (Spieler and Balota 1997).

The 636 stimuli were divided into six lists of 106 items each, comprising equal numbers of each stimulus type in each list. Each participant performed 1 of the 2 tasks on every stimulus, with the list order and task assignment counterbalanced across participants.

---

[8] Supplemental material B.3 and B.4 show the lexical property data presented in Table 3.1, but stratified by stimulus length.

Coding of response accuracy

Response accuracy for each participant for each item was determined by a single rater, during task performance. In cases of ambiguous response, the recording was used offline to determine correct coding.

For the regularize task, a pronunciation of a regular word was considered correct if it corresponded to the standard dictionary pronunciation. Pseudowords (e.g. /krot/ for "crote") were considered correct if their pronunciations followed typical spelling-to-sound correspondences, or rhymed with those of similarly spelled real words. A pronunciation of an irregular word was considered correct if it obeyed a phonologically legitimate alternative pronunciations of its rime (e.g., pronouncing "have" as /hev/). In some cases, more than one pronunciation of an irregular word was counted as correct. See Supplemental material B.2 for examples of irregular test stimuli along with possible correct responses.

All subjects achieved at least 65% accuracy in each task overall. Only the correct trials from each block were included. Items that were misidentified on the post-test for a given participant were excluded from the RT analyses. Because both groups performed both tasks with relatively high accuracy (and because of issues mentioned in Chapter 1 related to the multiplicity of error types and our inability to discriminate them confidently), analyses focused on reaction time effects. However, effects of accuracy were also examined and reported.

Calculation of response latency

Responses latencies were calculated from recorded files that had been purged for coughs, throat clearings, etc. using an in-house Matlab program (Matlab 7.80; R2009a; The MathWorks, Natick, MA) (Nelles et al. 2003). Reaction times were also spot-checked manually using Audacity, an open-source software for sound editing (http://audacity.sourceforge.net/).

The mean and standard deviation were calculated on an individual subject basis across all correct trials of both tasks, excluding raw reaction times < 200 ms or > 3 standard deviations

above a subject's mean as probable outliers (on average, 9.9 out of 636 trials were removed per subject). Individual differences in overall processing speed were controlled for by converting each response latency to a z-score based on each participant's mean reaction time and standard deviation (Faust et al. 1999). Repeated-measures analyses of variance (ANOVAs) were carried out at the subject- (F_1) and item- (F_2) levels to examine effects of age group, task and lexical variables (including lexicality, word regularity and word frequency). Appropriate post hoc analyses clarified the results of the planned comparisons.

Results

Raw RTs and accuracy rates are presented in Table 3.2[9]. Unless otherwise stated, response latency analyses are presented for the z-transformed RTs only. In general, the pattern of results was the same when analyses were computed on the untransformed RTs.

---

[9] Supplemental tables D.5 and D.6 show the raw reaction times and accuracy rates for the short and long items, respectively, considered separately.

Table 3.2 Mean raw reaction times (ms) and accuracy rates

Stimulus type	Read				Regularize			
	Adults		Children		Adults		Children	
	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)
High frequency, regular	678 (117)	1.000 (.000)	824 (125)	0.995 (.013)	1258 (277)	0.994 (.014)	1140 (209)	0.985 (.024)
Low frequency, regular	691 (124)	0.999 (.004)	835 (128)	0.984 (.031)	1205 (272)	0.994 (.018)	1105 (186)	0.980 (.026)
High frequency, irregular	704 (119)	0.996 (.008)	851 (124)	0.966 (.044)	1274 (267)	0.883 (.106)	1196 (223)	0.771 (.105)
Low frequency, irregular	746 (127)	0.979 (.023)	908 (155)	0.850 (.130)	1269 (264)	0.897 (.068)	1180 (232)	0.803 (.082)
Pseudo-words	841 (199)	0.992 (.012)	1004 (198)	0.926 (.107)	1126 (277)	0.988 (.024)	1135 (209)	0.948 (.058)

Children were faster than adults on the regularize task

Our first overarching question was whether children would show a larger, smaller or comparable reaction time cost compared to young adults, when instructed to direct attention entirely to phonological processing during the regularize task. As discussed in the Introduction, a reasonable case could be argued for a prediction in either direction, since 8-10 year old children likely have both weaker lexical route representations than adults (predicting, in simplified terms, a speed advantage) and relatively inferior inhibitory control (predicting, in simplified terms, a speed disadvantage). Thus the first analyses examined the effect of task (across all stimuli) and its modulation by age group.

Analyses on the z-transformed reaction times including the factors of task and age group showed a main effect of task ($F_1(1,62) = 332.55$, $p < .001$; $F_2(1,635) = 3001.21$, $p < .001$), an effect of age group that was significant by subjects only ($F_1(1,42) = 3.192$, $p = .003$; $F_2(1,635) = 0.110$, $p = .740$) and, more importantly, an interaction of task and age group ($F_1(1,61) = 17.93$, $p < .001$; $F_2(1,635) = 580.79$, $p < .001$; see Figure 3.1). This significant task x age group interaction was obtained for untransformed reaction times as well ($F_1(1,61) = 16.07$, $p < .001$; $F_2(1,635) = 668.89$, $p < .001$). The interaction arose because adults were 150 ms faster than children on the read task (zRT $t_1(61) = 4.03$; $p < .001$; $t_2(635) = 437.57$; $p < .001$) but 55 ms slower than children on the regularize task (zRT $t_1(61) = 4.41$; $p < .001$; $t_2(635) = 174.72$; $p < .001$).

Figure 3.1 Significant interaction of task x age group

RTs shown as *z*-transformed. Error bars (for all figures in this chapter) show +/- 1 standard error of the mean (SEM). Because the statistical analyses were computed as repeated-measures ANOVAs, the error bars may be used only to assess group variability; they cannot be used to infer statistical significance. Asterisks indicate significant main effects and interactions, with ***$p < .001$.
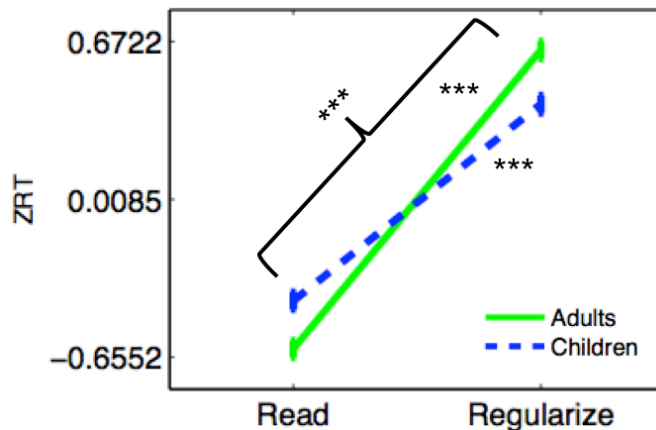
Analyses of the accuracy data including the factors of task and age group also showed main effects of task ($F_1(1,62) = 55.94$, $p < .001$; $F_2(1,635) = 32.04$, $p < .001$) and age group ($F_1(1,37^*) = 5.328$, $p < .001$; $F_2(1,635) = 436.149$, $p < .001$), as well as an interaction of task and age group that reached significance at the item-level only ($F_1(1,61) = 0.11$, $p = .741$; $F_2(1,635) = 7.71$, $p < .010$). Both children and adults were more accurate on the read task than the regularize task (children +3.0%; $F_1(1,30) = 18.77$, $p < .001$; $F_2(1,635) = 7.10$, $p < .010$; adults +3.3; $F_1(1,31) = 44.05$, $p < .001$; $F_2(1,635) = 88.88$, $p < .001$). At 0.3%, the age group difference was clearly small.

Although it is robust and remarkable, the task x age group interaction for reaction time is difficult to interpret, since several potentially interactive age group differences may contribute, including the strength of lexical representations, the efficacy of selective attention, grapheme-to-phoneme conversion facility and the nature of a post-lexical response check. We therefore turned to our second question, to ask whether children and adults show differences in the way that stimulus characteristics interact with task. The logic was that stimulus modulation of the task x age group interaction may be useful for distinguishing developmental differences in

pathway control from developmental differences in response selection, in particular because any task interaction with regularity as well as any task effect for pseudowords would be difficult to attribute to lexical pathway interference alone (as suggested by Balota et al. 2000).

The next analyses examined whether frequency and/or lexicality would modulate the robust reaction time interaction of task and age group. First, frequency, the more subtle of the 2 properties, was addressed. Following Balota et al., the analyses of frequency and lexicality included only regular words and pseudowords. As those authors pointed out, both regular words and pseudowords generate the same pronunciations for the 2 tasks, so their task comparison is free of potentially confounding articulatory effects.

The frequency effect for regular words in the regularize task showed a weak age group effect

Treating frequency as high vs. low did not reveal an age group difference in the frequency effect for regular words in the regularize task

Analyses of zRT were conducted using task, frequency and age group as factors. The main effects of task and age group were discussed above with regard to the task x age group analyses. Consistent with Balota and colleagues, there was a reliable task x frequency interaction ($F_1(1,62) = 30.98$, $p < .001$; $F_2(1,190) = 15.62$, $p < .001$). The task x frequency interaction was obtained in each age group considered separately, though it was statistically less reliable in the children (adults: $F_1(1,31) = 30.55$, $p < .001$; $F_2(1,190) = 16.97$, $p < .001$; children: $F_1(1,30) = 7.31$, $p = .011$; $F_2(1,190) = 5.70$, $p = .018$) (see Figure 3.2). The interaction of task and frequency across the age groups was driven by a positive word frequency effect in the read task that was significant at the subject-, but not item-, level ($F_1(1,62) = 5.30$; $p = .025$; $t_2(190) = 1.60$; $p = .111$) and a significant negative word frequency effect in the regularize task ($F_1(1,62) = 31.67$; $p < .001$; $t_2(190) = 2.80$; $p = .006$).

Figure 3.2 Significant interaction of task x frequency in each age group
The interaction of task x frequency (computed using regular words only) was significant in each age group separately, with no significant 3-way interaction by rmANOVA. *z*-transformed RTs are plotted, and raw RT differences between high and low frequency are indicated for each task for each age group. Read task shown in black dotted lines, regularize task shown in red solid lines. *p* values depicted for both the subject- and item- analyses (subject-/item-), where ***p < .001; **p < .01; *p < .05.
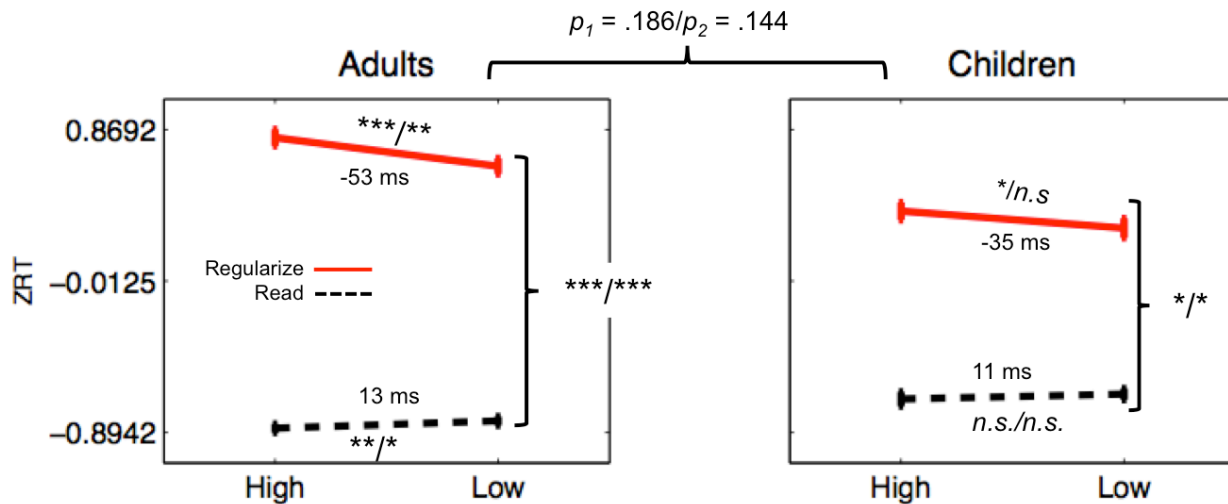
The 3-way interaction of task, frequency and age group was not significant by rmANOVA ($F_1(1,61)$ = 1.79, p = .186; $F_2(1,190)$ = 2.16, p = .144). However, and as shown in Figure 3.2, *post hoc* evaluations of the task- and age group- specific frequency effects suggested different patterns in the 2 groups. Specifically, the word frequency effect reached significance for the adults in *both* tasks (read $F_1(1,31)$ = 7.91, p = .008; $t_2(190)$ = 2.06; p = .041; regularize $F_1(1,31)$ = 22.17, p < .001; $t_2(190)$ = 3.12; p = .002), but for children only in the regularize task, and only at the subject- level (read $F_1(1,30)$ = 1.07, p = .310; $t_2(190)$ = 1.04; p = .302; regularize $F_1(1,31)$ = 7.31, p = .011; $t_2(190)$ = 1.67, p = .096). Stated in terms of effect sizes, and as shown on the plots, in adults the word frequency effect changed from +13 ms for the read task to -53 ms for

the regularize task, thus reversing by 66 ms[10]. In children, the task-induced word frequency effect was smaller in magnitude by 30%, changing 46 ms from +11 ms for read to -35 ms for regularize.

Treating frequency as a continuous variable illuminated an age group difference in the frequency effect for regular words in the regularize task

As just described, the categorical analysis (rmANOVA) did not yield a significant 3-way interaction of task x frequency x age group for the regular words (F_1 = 1.79; F_2 = 2.16). However, children showed a change in the word frequency effect (from the read task to the regularize task) that was 30% smaller in magnitude than that observed in the adults, and *post hoc* evaluations of the frequency effects obtained in each age group for each task suggested that the groups were not the same. Given our interest in determining whether developmental differences exist in terms of both pathway control and response checking, it was imperative to pursue the possibility that there existed even a weak interaction of task, regularity and frequency with age group.

We therefore next conducted a set of *post hoc* correlation analyses to explore the possibility that age group differences in the relationship between task and frequency for regular words exist, but are small enough so as not to emerge when frequency is treated as a dichotomous variable. As described in Chapter 2, further support for the appropriateness of conducting a correlation analysis of the effect of frequency is the fact that the word stimuli were selected intentionally to sample a continuous range of the log-transformed Zeno et al. Standard Frequency Index (SFI) (1995) (see Supplemental figure B.8), and the original choice to dichotomize the stimuli by a median split was a matter of convenience, rather than a theoretical commitment to thinking about frequency as a simple construct of high vs. low.

---

[10] The task-induced word frequency effect reversal in the adults in the present study is quite similar to what was reported in the Balota et al. report (from +26 ms for read to -43 ms for regularize, reversing 69 ms).

Frequency was coded as a continuous variable, and correlations were computed as described in Chapter 2 with regard to the interaction of task x frequency x regularity in adults. Item-level correlations between group-average zRT and log frequency were calculated using all 384 word stimuli (n = 192 each of regular and irregular words). Each item was entered into four bivariate correlations, one for each task (read and regularize) for each age group (children and adults). Regular and irregular words were treated separately.

As shown in Figure 3.3, the correlations revealed age group differences in the modulation, by task and by regularity, of the relationship between frequency and reaction time. As reported in Chapter 2 (Figure 2.5), frequency facilitated read reaction times for both regular (R = -.176; *p* = .015) and irregular words (R = -.312; *p* < .001) for the adults. Also as previously stated, frequency significantly *inhibited* regularize reaction times for the adults for regular words (R = .242; *p* = .001), but not irregular words (R = -.046; *p* = .524), consistent with the 3-way interaction first reported by Balota and colleagues (2000). Figure 3.3 shows that for the children, on the other hand, the pattern of task, frequency and regularity was clearly different. Frequency facilitated read reaction times for the irregular (R = -.400; *p* < .001), but not the regular words (R = -.099; *p* = .171). Critically, and in contrast to what was observed in the adults, the children showed a null effect of frequency on regularize task reaction time for both regular and irregular words (regular R = .096; *p* = .184; irregular R = .051; *p* = .482).

The regularity- and age group- dependence of the relationship between frequency and task, as revealed by the *post hoc* correlations, does not constitute a direct test of difference, and should therefore be interpreted cautiously. To assess further the confidence with which we should interpret the correlation findings, we next computed 4-way interaction between frequency, regularity, task and age group by rmANOVA. That interaction reached significance at the item-level and approached significance at the subject- level (F_1(1,61) = 3.09, *p* = .084; F_2(1,380) = 5.69, *p* = .018).

Figure 3.3 Item-level age group- and regularity-specific correlations between frequency and reaction time for both tasks

     In the regularize task, only the adults showed a significant negative correlation between frequency and reaction time, and this was only true for regular words. Reaction times for each word stimulus are plotted along the y-axes, averaged either across adults (top panel) or children (bottom panel), separately for the read task (left panel) and the regularize task (right panel). Frequency is shown along the x-axes, as the standard frequency index (a log-transformed count per million) from the Zeno et al. norms (Zeno et al. 1995). Asterisks and solid lines indicate a significant linear relationship, where ***$p \leq .001$; *$p \leq .05$. Dotted lines indicate the absence a significant linear fit. Regular words are shown as orange diamonds, while irregular words are plotted as periwinkle circles. The x- and y-axis scales are the same for all 4 plots.
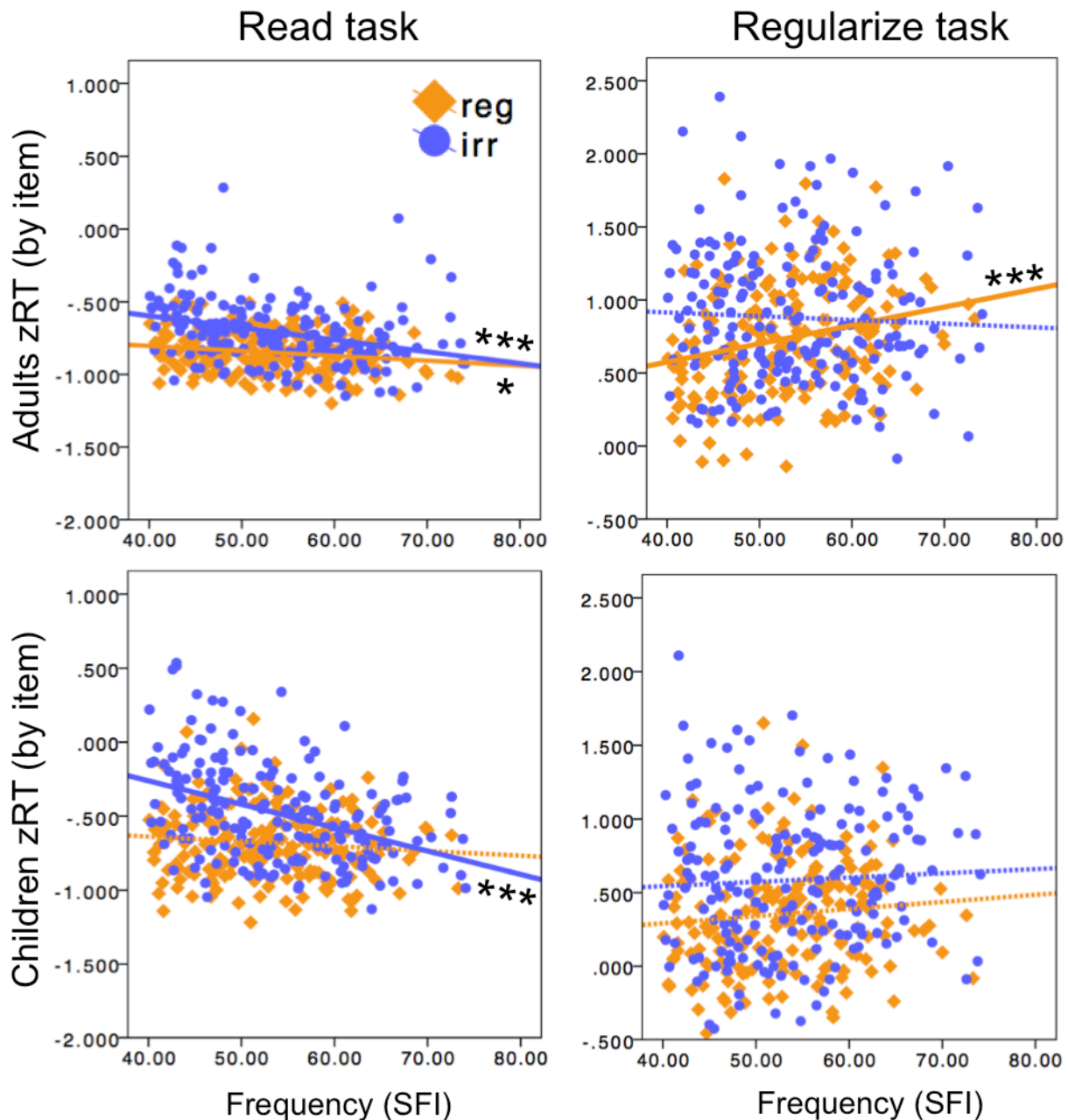
Categorical analyses of accuracy for task, frequency and age group were equivocal

Returning to the treatment of frequency as a categorical variable using rmANOVAs, analyses of the accuracy data were also computed using the factors of task, frequency and age group. Compared to the reaction time results, effects of the 3 factors of interest were relatively subtle, and the F_1 and F_2 analyses did not converge. First, the interaction of task x frequency for accuracy reached significance at the item- level ($F_2(1,190) = 4.77$, $p = .030$) but not the subject-level ($F_1 < 1$). The task x frequency interaction was further modulated by age group, again at the item- ($F_2(1,190) = 4.25$, $p = .041$), but not the subject- level ($F_1 < 1$). The 3-way interaction was driven by a significant item-level interaction of task and frequency in the children ($F_1 < 1$; $F_2(1,190) = 5.27$, $p = .023$) that was absent in the adults (both $Fs < 1$).

Given the weak but suggestive evidence that, for reaction time, age group interacts with task, frequency and regularity, the next analyses asked whether the 2 age groups would also show differences in the interaction between task and lexicality. As discussed in the Introduction, a task x age group effect for pseudowords, in particular, may corroborate the notion that the groups differ with regard to the contribution of the response check step, since pseudowords do not have lexical representations.

Lexicality modulated the interaction of task and age group, but adults were relatively slower for both words and pseudowords

Analyses of zRT were next conducted using task, lexicality and age group as factors. Consistent with the Balota et al. study, the interaction of task and lexicality was significant ($F_1(1,62) = 265.30$, $p < .001$; $F_2(1,442) = 632.61$, $p < .001$). More importantly, the relationship between task and lexicality was robustly modulated by age group ($F_1(1,61) = 9.40$, $p < .005$; $F_2(1,442) = 34.83$, $p < .001$; see Figure 3.4). The 3-way interaction was driven by a reliable age group difference in the effect of lexicality on the regularize task ($F_1(1,61) = 31.87$, $p < .001$;

125

$F_2(1,442) = 121.59$, $p < .001$). In particular, children showed a null lexicality effect for the regularize task (+12 ms; $F_1(1,30) = 0.37$, $p = .550$; $t_2(442) = 0.84$, $p = .404$), while adults showed a strong *reversed* lexicality effect for the regularize task (-106 ms; $F_1(1,31) = 49.07$, $p < .001$; $t_2(442) = 9.41$, $p < .001$). On the other hand, both groups showed large, positive, reliable lexicality effects for the read task (adults = +156 ms: $F_1(1,31) = 238.71$, $p < .001$; $t_2(351*) = 20.77$, $p < .001$ children = +175 ms: $F_1(1,30) = 59.57$, $p < .001$; $t_2(418*) = 19.00$, $p < .001$). The lexicality effects in the read task were not reliably different between age groups at the subject-level ($F_1(1,61) = 2.17$, $p = .146$), though the 2-way interaction did reach significance at the item-level ($F_2(1,442) = 11.69$, $p < .001$), reflecting a larger lexicality effect in the read task for the children as compared to the adults.

A complementary way of decomposing the significant 3-way interaction is to examine the interaction of task and age group for words and pseudowords separately. For regular words, both groups showed a main effect of task (adults $F_1(1,31) = 755.40$, $p < .001$; $t_2(191) = 59.30$, $p < .001$; children $F_1(1,30) = 163.09$, $p < .001$; $t_2(191) = 39.12$, $p < .001$), but the effect of task was significantly greater in the adults than the children ($F_1(1,61) = 31.30$, $p < .001$; $F_2(1,191) = 344.96$, $p < .001$). Similarly, for pseudowords, both groups showed a main effect of task (adults $F_1(1,31) = 59.21$, $p < .001$; $t_2(251) = 40.14$, $p < .001$; children $F_1(1,30) = 34.93$, $p < .001$; $t_2(251) = 20.12$, $p < .001$), and, as with regular words, the effect of task for pseudowords was significantly greater in the adults than the children ($F_1(1,61) = 5.81$, $p = .019$; $F_2(1,251) = 119.34$, $p < .001$).

Figure 3.4 Significant interaction of task x lexicality x age group

Age group reliably modulated the interaction of task x lexicality (computed using regular words and pseudowords). Still, for both words and pseudowords, the task reaction time cost (distance from the black line to the red line) was, on average, greater in the adults than the children. *z*-transformed RTs shown. Read task shown in black dotted lines, regularize task shown in red solid lines. *p* values depicted for both the subject- and item- analyses (subject-/item-), where ***p < .001; **p < .01.
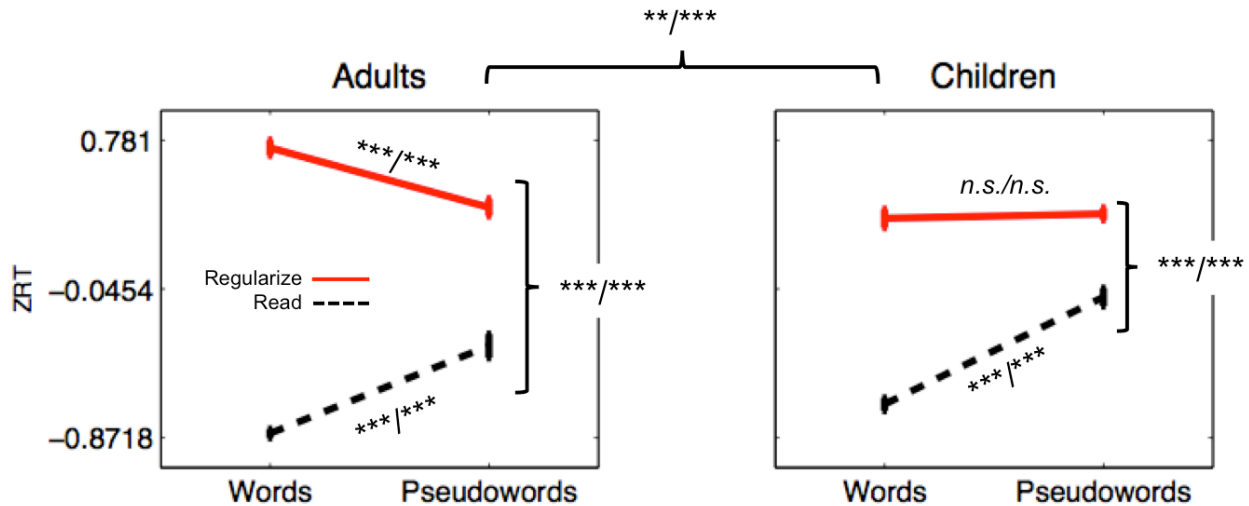
Turning to the accuracy data, the interaction of task and lexicality was significant ($F_1(1,62) = 61.54$, $p = .016$; $F_2(1,442) = 19.67$, $p < .001$). As with reaction time, the relationship between task and lexicality for accuracy was further modulated by age group ($F_1(1,61) = 4.53$, $p = .037$; $F_2(1,442) = 12.60$, $p < .001$), however this is in part due to a ceiling effect for the adults. Children were 6.4% more accurate for regular words (99.0%) than for pseudowords (92.6%) in the read task, and just 3.4% more accurate for regular words (98.2%) than for pseudowords (94.8%) in the regularize task. Adults, on the other hand, performed close to ceiling for both regular words and pseudowords in both tasks (>98% or regularize, >99% for read).

The fact that the task reaction time cost was most prominent for words for the adults (compare distances between black and red lines in Figure 3.4 for each task for each age group, or refer to Table 3.2 for raw data) is consistent with the notion that adults encounter particular difficulty inhibiting the lexical pathway during the regularize task (Balota et al. 2000). However,

the task effect was shown to be significantly more pronounced in adults than in children for

pseudowords as well. Specifically, as highlighted in Figure 3.5, even though adults were 164 ms

faster than the children to *read* pseudowords (raw RT $t_1(61) = 477.37$; $p < .001$; $t_2(251) = 3.26$; p

$= .002$), the 2 groups showed a non-significant 9 ms RT difference for pseudowords for the

regularize task (raw RT[11] $t_1(61) = 0.14$; $p = .892$; $t_2(251) = 0.01$; $p = .916$). Stated differently, the

adults were 285 ms slower to regularize pseudowords than to read them, while the comparable

task difference in children was significantly smaller, by more than half (131 ms).

Figure 3.5 Interaction of task x age group for pseudowords only
     Reaction times shown as raw RTs, to emphasize that the age group differences were absolute, not just relative. Red text and brackets emphasize the age group effect for each task (read task >> regularize task), while black text and brackets emphasizes the task effect in each age group (adults > children). *p* values depicted for both the subject- and item-analyses (subject-/item-), where ***p < .001; **p < .01.

---
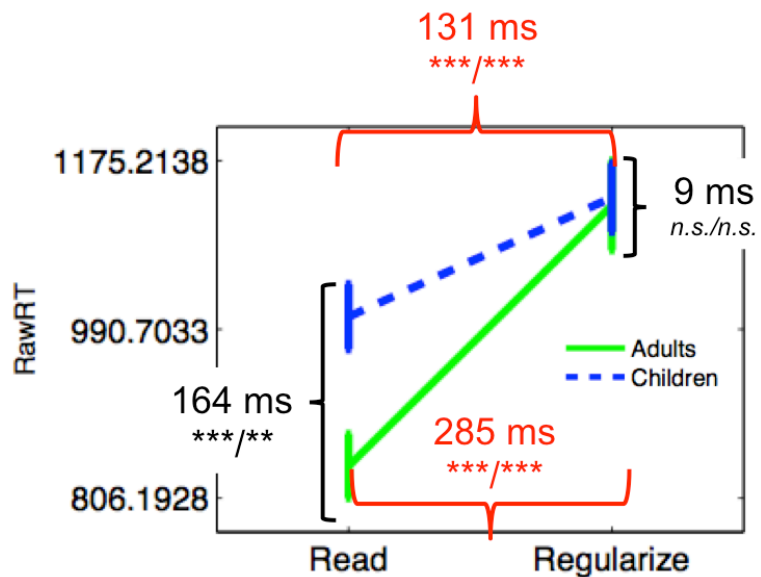
[11] Analyses showed the same statistical results for the zRTs, so raw RTs are reported to emphasize that the age group interaction is absolute, not just relative.

128

Discussion

This study compared performance on a read and a regularize task in young adults and school-age children, in order to investigate the interaction of attentional control and reading and, in particular, how that interaction may be different for early and experienced readers. Extending a previous study (Balota et al. 2000) as well as the work in Chapter 2, we asked 2 primary questions in this chapter. First, we asked whether 8-10 year old children would show larger, smaller or comparable reaction time costs, compared to young adults, for performing the regularize task. Second, we asked whether children and adults would show differences in the interactions of particular stimulus features with task, potentially discriminating the relative contributions of developmental differences in pathway control from developmental differences in response selection. Framing our study to ask these 2 primary questions allowed us to pursue twin objectives of re-assessing the evidence for the 2-mechanism model (following both the original Balota et al. study and the data presented in Chapter 2) while simultaneously extending those previous studies by exploring differences between 8-10 year old children and young adults.

Together, the results of this chapter provide converging evidence in support of the Balota et al. suggestion of a 2-mechanism explanation for attentional control of the regularize task, with roles for attentional modulation at both the level of pathway control and response selection, thereby also corroborating the findings presented in Chapter 2. The results furthermore suggest that the development of reading likely involves changes in *both* of the proposed mechanisms of attentional control over reading, although this latter conclusion should be corroborated by attempts both to replicate the behavioral data and to compare the 2 age groups using neuroimaging (see Chapter 4, future directions).

A reading task was identified in which children were faster than young adults

The first set of analyses revealed a robust task x age group crossover for reaction time, driven by the adults being faster than the children on the read task, as expected, and the children being faster than the adults on the regularize task, which was not expected. It is frankly surprising to identify a task in which children were reliably faster (though less accurate) than adults, and this observation alone is indeed one of the most striking and noteworthy of this study. Indeed, children are slower and less accurate than young adults (Elliott 1970; Wickens 1974) on just about any cognitive task (Hale 1990; Kail 1991), and processing speed, measured as reaction time, has been shown to decrease robustly from childhood through adolescence to adulthood (Kail and Miller 2006). Despite the overwhelming literature precedent, the reaction time advantage for children over adults in the regularize task was large and highly reliable, and it was separately evident in sub-analyses considering regular words, irregular words and pseudowords separately.

Since children are generally slower than adults to perform a task, what could account for our finding that the children were, on average, 55 ms faster than the adults on the regularize task? Unfortunately, a clear interpretation of a reaction time advantage in favor of children in this case is difficult, since several potentially competing and interacting factors that may differentiate children from adults must be considered, including the ability to attend selectively to one process while inhibiting another (i.e., exert pathway control); the ability to map graphemes onto phonemes; the strength of lexical route representations; and the quality and duration of the response checking. Because of the large number of degrees of freedom, and because the task x age group interaction collapsed across all stimulus types, additional analyses were conducted to address our second primary question. Specifically, we investigated whether stimulus features (frequency, regularity and lexicality) would interact with the task x age group interaction, thereby

illuminating the source(s) of the faster overall regularize performance by the children as compared to the adults.

An age group difference in the frequency effect for regular words in the regularize task points to a developmental difference in response checking

A 3-way rmANOVA suggested that frequency did not reliably modulate the observed task x age group interaction when treated as a dichotomous variable. Indeed, the categorical analysis approach revealed a reliable interaction of frequency x task for regular words that reached significance in each age group separately. However, several factors led us to perform a *post hoc* analysis in which frequency was treated as a continuous, rather than a dichotomous, variable, to test the possibility that *weak* age group differences in the interplay between task, regularity, frequency and age group exist. Conceptually, we were wary of incurring a Type II error in assessing for age group differences in this particular analysis, since, as described earlier, this higher-level interaction is one of the key ways to assess for the contribution of a post-lexical response check to regularize task performance (Balota et al. 2000). Furthermore, the magnitudes and statistical reliabilities of the frequency effects for each task for each age group hinted at underlying developmental differences. Methodologically, the stimulus set was composed of words that spanned a frequency continuum and were subjected to a median split, meaning that a categorical analysis approach to effects of frequency was likely, *a priori*, be less sensitive to weak effects than a continuous analysis approach. Lastly, a suggestive (significant for F_2 and trend-level for F_1) *post hoc* 4-way interaction of the variables of interest was obtained, boosting our confidence in the age group differences revealed by the correlations.

Recognizing, then, both the *post hoc* nature of the correlations and the fact that, as a set, they do not constitute a direct, statistical test of significance, we were nonetheless intrigued to observe a clearly different pattern in the children as compared to the adults, as depicted in Figure 3.3. In particular, with regard to the specificity of the frequency effect reversal for the

regularize task (it was only obtained for regular words, and only in the adults), the correlation results make it difficult to reject the idea that there is a real, if relatively weak, difference in the way that 8-10 year old children and adults process regular words in the regularize task. Cautiously, we interpret the correlations as evidence that the adults incur a greater slow-down, than do children, from response checking during the regularize task. In interpreting the effect of regularity in this manner, we are concurring with the analysis presented by Balota and colleagues with respect to interference from the lexical pathway: "there should be no influence of regularity (because this effect presumably is due to competition between the two routes)" (Balota et al. 2000).

Adults were relatively slower, compared to children, to regularize than to read both words and pseudowords

Given the evidence suggesting, perhaps speculatively, that the word frequency effect reversal for regular words in the regularize task was weaker in the children than in the adults, a third set of analyses examined whether lexicality would modulate the task x age group effect. Because lexicality is essentially a strong frequency manipulation, age group effects of lexicality should, if anything, be larger than age group effects of frequency. Furthermore, examining the effect of age group on the task difference separately for words and pseudowords provides a way to tease apart age group differences due to pathway control (presumably applicable to words) from those arising from the response check step (potentially applicable to both words and pseudowords).

Analyses of task, lexicality and age group revealed a significant 3-way interaction of these variables for reaction time. While both children and adults showed a significant interaction of task x lexicality, the nature of that interaction was qualitatively and statistically different in the 2 age groups. The age group difference was significant whether analyses were computed using z-transformed reaction times or raw reaction times, i.e. whether each individual's reaction times

132

were normalized or not. Specifically, adults showed a significant and reversed lexicality effect for the regularize task, while lexicality did not appreciably influence regularize task reaction times for the children.

Within the robust task x lexicality x age group interaction, both the words and the pseudowords showed separate significant interactions of task x age group. The fact that adults are relatively slower than children to regularize both words *and* pseudowords is a crucial observation, as will now be discussed.

Slower regularize performance for words in the adults suggests an age group difference in pathway control

The large task x age group interaction for words likely derives, at least in part, from greater lexical route interference in the adults than the children, i.e., from age group differences in pathway control. As discussed earlier, the adult cohort has a greater cumulative exposure to written word forms than the child cohort (Stanovich 1993; Share and Stanovich 1995), which presumably results in either lower thresholds of activation or higher resting activations for individual lexical entries in the adults than the children. On the other hand, adults ought to be better able than children to exert the appropriate bias signal at the level of pathway control, consistent with the general observation that the deployment of selective attention is still very much developing in 8-10 year olds (e.g., Rueda et al. 2004; Luna et al. 2004; Zelazo et al. 2004; Davidson et al. 2006). One interpretation of the observation that the task x age group interaction for words was larger than the corresponding interaction for pseudowords is that the dominance of the lexical route representations in the adults overwhelms the adults' presumably superior ability to attend selectively to one component process. These results are consistent with the notion that adults experience greater lexical route interference than the children when performing the regularize task. Stated differently, 8-10 year old children may be better able than adults to exert pathway control over the component processes of reading, presumably not

because of their superior selective attention, but because their lexical representations are so much less well-established. (See Supplemental analysis D.2 for an examination of how the notion of pathway control in the context of the regularize task related to Stroop performance in these participants.)

Slower regularize performance for pseudowords in the adults corroborates an age group difference in response checking

In contrast to the situation with words, overwhelmingly strong lexical representations (resulting in poor pathway control) seem insufficient to account for the relative pseudoword slowing in the regularize task in the adults compared to the children (highlighted in Figure 3.5). If pathway control of prepotent lexical representations were the only relevant attentional factor, pseudowords should arguably be regularized *faster* than they are read. That is, emphasis on the phonological pathway during the regularize task should be, even if it cannot be achieved completely, relatively greater than emphasis on the phonological pathway during the read task. This prediction follows from the notion that the read task presumably entails a relative unweighting of the phonological pathway, in favor of more rapid lexical route processing (recall that the stimuli appear in mixed lists). Not only would a pure pathway control account predict faster reaction times for pseudowords for regularize than for read – for both age groups – but it would also by extension predict that any reaction time advantage that the adults show for pseudowords in the read task would be preserved, or perhaps magnified[12], in the regularize task. Instead, we report that adults were 164 ms faster to read pseudowords than children, but only 9 ms faster to regularize the same kinds of pseudowords.

---

[12] A magnification of the speed advantage for pseudowords for adults (compared to children) in the regularize task (compared to the read task) would follow from the fact that the adults would be expected to bias lexical route processing (rather than phonological route processing) quite heavily, and certainly moreso than the children during the read task.

134

It is our contention that the task x age group interaction for pseudowords is not only noteworthy in and of itself, but is also compelling evidence for an age group difference in the response check step. In this way, the pseudoword task x age group effect corroborates the observations from the correlation analyses involving task, frequency, regularity and age group. Nonetheless, and especially in light of evidence suggesting the plausibility of fine-grain dynamic attentional control over nonword naming at the trial-by-trial level (Reynolds et al. 2012), it is reasonable to speculate that at least some of the slowdown for the pseudowords in the regularize task could arise from feedforward activation of the lexical route (Coltheart et al. 2001) for those particularly "word-like" pseudowords. In such a scenario, the larger task effect for pseudowords in adults than in children would result from a greater degree of partial (or erroneous) lexical route activation by pseudowords in the adults. Although this conceptualization is not the one that we favor, it raises the possibility that the task effect for pseudowords could be accommodated by a single-mechanism age group difference in terms of control over the lexical pathway. One way that the pseudoword slowdown for the regularize task could be further interrogated in terms of a pathway control explanation is to examine the ways in which bigram frequency and orthographic neighborhood size (properties related to "wordness") co-vary with pseudoword regularize reaction time. At this point, we think the data are best accommodated by a 2-mechanism account.

Supplemental analysis D.7 presents an exploratory subject-level correlation analysis that further suggests the special nature of the task x age group effect for pseudowords (as compared to the task x age group effect for words).

Suggestion of one interpretive framework

Related to the consideration of whether the process of sounding out an item (as in the regularize task) is *necessarily* heavily influenced by that item's lexical representation, Reynolds and Besner presented some interested findings suggesting that lexical representations need not

automatically contribute to the computation of phonology (2006). Combining a psychological refractory period paradigm with a speeded naming task, those authors showed that orthographic neighborhood size modulated pseudoword naming reaction times in a manner that was additive rather than interactive with stimulus onset asynchrony (SOA)[13], implying that phonological decoding requires central processing resources and is thereby not automatic/obligatory. Another experiment showed that the effect of repetition on speeded naming of low-frequency exception words did interact with SOA, suggesting that the earliest components of lexical processing do not require attention and may therefore fulfill at least one of the primary criteria for automaticity. Together with several additional experiments, the data presented by Reynolds and Besner suggest a distinction between an early, automatic component of lexical processing (orthographic-lexical) and a later, attention-requiring component (phonological decoding) (2006). Such a conceptualization departs from the widely-held view that visual word recognition writ large, and phonological decoding in particular, proceeds largely automatically (e.g., LaBerge and Samuels 1974; Perfetti 1992; Booth et al. 1999), and may prove useful for thinking about the results of the present study.

Specifically, the Reynolds and Besner model leaves unanswered the question of whether the stage labeled "phoneme buffer" is subject to attentional modulation in adults. Furthermore, because the Reynolds and Besner model derives from data on single-word reading in adults, it is interesting to ponder how the observations of the current study, which includes both a task variant (regularize) and a developmental component, could be useful for fleshing out their model. One way to integrate our results with theirs is to assert that our data underscore the importance of the "phoneme buffer" stage for an attentionally demanding variant of speeded naming, such as the regularize task. Additionally, our results suggest that the

---

[13] Varying the stimulus onset asynchrony (SOA) between 2 consecutively presented tasks allows for an investigation of the extent to which the tasks rely on presumably shared (and finite) attentional resources.

importance of attentional modulation of the "phoneme buffer" stage is greater in adults than in children, i.e., that there are important developmental effects at this relatively late stage of lexical processing. An interpretation of this sort accommodates the slower overall regularize performance in the adults as compared to the children, as well as the 2 critical age group effects of stimulus (i.e., the *post hoc* frequency effect reversal for regular words for the regularize task in adults, and the task x age group interaction for pseudowords). More broadly speaking, the assertion of reduced significance of response checking in the children, compared to the adults, also reverberates with the notion that children enjoy a learning advantage for language partly by virtue of freedom from top-down control over the component processes (Chrysikou et al. 2011).

Balota et al. (2000) discussed previously the need to go beyond a static view of lexical processing and consider the flexibility of the system in terms of pathway attenuation and accentuation The present developmental results reassert that claim, and argue for an expanded version of that suggestion. We propose that not only the relative weighting of the processing pathways but also the step of response selection may be modulated by attentional control, depending on task demands *and* age/level of proficiency. Future studies will attempt to replicate these behavioral data, and will also use task-based functional MRI (fMRI) to investigate the neural correlates of the developmental differences presented here. For example, fMRI may be able to identify age group differences in the activity of sets of brain regions associated with pathway control on the one hand, and response selection on the other (c.f., Chapter 2). Possible approaches to the planned developmental imaging study are discussed in Chapter 4.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

In this final chapter, I will begin by summarizing the main observations of Chapters 2 and 3. In the context of two ways in which the reading literature is arguably somewhat conceptually stagnant, this thesis work may represent something of a rogue approach to thinking about reading and its development. I will next explore 2 additional observations about the data that were not directly addressed by either of the data chapters' discussion sections.

The remainder of the chapter will then explicate a series of ideas for future studies that may serve to strengthen (or, equally usefully, call into question) the notions that (1) attentional control interacts with reading-related tasks in at least 2 distinct ways, and that (2) both of these interacting mechanisms are immature in early readers. Some of proposed future analyses clearly indulge in more speculation than others, and I will end that part of this chapter by describing the most proximally planned follow-up, a developmental imaging comparison. I will also briefly endorse the importance of pursuing any of the suggested future experiments in the first place. I will finally conclude by pointing out that a 2-mechanism account of controlled lexical processing – like the one proposed here, combining pathway control and a response check – finds itself in decent company in the literature.

Summary and significance

Chapters 2 and 3 presented a series of analyses of 2 single-word reading tasks, read (simple speeded naming) and regularize ("sound it out") (Balota et al. 2000), focusing on the ways in which stimulus characteristics interacted with task instructions (both chapters) and age group (Chapter 3). Considering the totality of the data – and acknowledging that not every single observation ought to be interpreted equally confidently – I would submit that this thesis furnishes solid evidence for 2 mechanisms of interaction between reading and attentional control. Support of this contention converges from a replication of the essential behavioral findings of a previous

study (Balota et al. 2000), in fact building upon those authors' idea of a 2-mechanism account, in combination with novel neuroimaging (Chapter 2) and developmental (Chapter 3) insights.

With the novel neuroimaging angle, we identified 2 sets of regions, including members of previously defined attentional control subsystems, that showed activity mapping onto the 2-mechanism model. Regions showing signals reflecting pathway control included members of bilateral Cingulo-opercular control network and bilateral Dorsal attention system, plus a few others (CO/DA+). Regions showing activity consistent with a role in response checking included several left-lateralized regions from the Frontal-parietal control network (LFP).

With the novel developmental comparison, we established that children were, overall, faster than the adults to perform the regularize task. We also observed age group differences in reaction time in several key interactions, including equal performance for the 2 groups to regularize pseudowords, as well as an intriguing correlation result suggesting a different handling of high frequency regular words in the adults as compared to the children. The entire pattern of age group differences suggested that both pathway control and response checking are only partially mature in early readers.

<div align="center">**Further observations**</div>

Neuropsychological test results can be leveraged to understand task performance

Both the behavioral and the imaging studies reported in this work included the collection of neuropsychological testing for all participants. Neuropsychological testing is conducted experimentally because the scores may illuminate individual differences in task performance, by extension providing insight into factors distinguishing task performance between groups. For example, in discussing the correlations between the lexicality effects in the read and the regularize tasks in Supplemental analysis C.1, I made the point that the observation of a within-subject correlation between a large lexicality reversal in the regularize task and a small positive

lexicality effect in the read task was incongruent with a pure pathway control account of the attentional control required for the regularize task. Specifically, I argued that if lexical pathway interference can entirely explain the slower performance on the regularize task as compared to the read task (minimizing a contribution from a post-lexical response check), then it should follow that those subjects who have strongest lexical representations should have both the largest negative lexicality effects in the regularize task *and* the largest positive lexicality effects in the read task.

One might alternatively take the position that the correlation reported in Supplemental analysis C.1 between small, positive lexicality effects in the read task and large, negative lexicality effects in the regularize task is not as surprising as I remarked it to be. For example, presume, as has been proposed (e.g., Pugh et al. 2001) that reading skill development entails relative decreases in reliance on phonological processing and concomitant increases in reliance on lexical processing. In this scenario, the most facile readers would be predicted to show the largest positive lexicality effects during the read task, reflecting a push/pull between decoding and whole-word recognition skills (Baron and Strawson 1976). Those very good readers might also be expected to be the most flexible in their ability to perform any lexical task, given their generally superior reading abilities, which would in turn lead to the prediction that they would show the smallest interference effects in the regularize tasks. While this set of priors is not the one that I would adopt in considering this question, it is a certainly a defensible position.

However, in contrast to the notion that good readers sacrifice phonological processing skills in exchange for gains in lexical route proficiency, several reports have shown that the best readers show the smallest effects of many lexical variables, implying that there is no tradeoff, during reading skill development, between use of phonological and orthographic information (e.g., Brown et al. 1994; Schilling et al. 1998; Seidenberg 1985; Chateau and Jared 2000; Yap et al. 2012). As explained in Supplemental material D.4, the best readers in the present data, by

140

converging measures, showed the smallest effects of most variables except, prominently, the lexicality reversal in the regularize task.

Specifically, among both children and adults, the fastest readers, in terms of raw reaction time for the read task, showed the smallest lexicality effects for the read task; for both groups, the correlation was large and very reliable. Relatedly, the fastest readers showed the largest negative lexicality effects in the regularize task, although this effect reached significance only in the children (and was trend-level in the adults). In addition, among the adults, the participants who were fastest on the regularize task (in terms of raw reaction time) showed the largest negative lexicality effects in the regularize task (this relationship was null in children). Lastly, higher Letter/Word ID scores (requiring sight-word recognition) correlated with both smaller lexicality effects in the read task and larger negative lexicality effects in the regularize task, and these relationships were significant in both children and adults. Together, these results suggest that the best readers (relative to their ages) were characterized by small positive effects of lexicality in the read task and large negative effects of lexicality in the regularize task. The pattern observed at the level of individual participants was recapitulated at the group level, whereby the better readers (in absolute terms, the adults) showed larger interference effects in the regularize task than did the worse readers (in absolute terms, the children).

Relatedly, although existing evidence is rather equivocal with regard to whether the lexicality effect in speeded naming increases or decreases across skill development (Defior et al. 1996; Sprenger-Charolles et al. 1998; Fernandes et al. 2008), we reported in Chapter 3 that, if anything, children showed larger positive lexicality effects in the read task than adults (the effect was significant at the items, but not the subject-, level). Larger lexicality effects in the read task for the children as compared to the adults corroborates the individual neuropsychological and task performance correlations in suggesting that good readers show small effects of lexical

141

variables in speeded naming tasks, regardless of whether those variables seem most likely to map onto phonological or lexical pathway processing.

The children were faster than the adults to regularize, but they made more errors

One concern that could be raised about the data presented in Chapter 3 is that an age group difference in the tradeoff between responding quickly and responding correctly may have contributed to at least a portion of the observed task x age group reaction time effect. There are several reasons that I think it is reasonable to discount any serious suspicion that developmental differences in the speed-accuracy tradeoff contributed significantly to the overall reaction time patterns.

First, let us recall the observations. The astonishing finding from Chapter 3 was that children were *faster* than the adults to regularize. However, they also, on average, made more errors than did the adults; collapsed across all stimulus types, children were 5% less accurate than the adults on the regularize task (96% vs. 91%). Note, also, that the average accuracy of 91% in the children is slightly misleading, as it does not account for the data from the 3 children who were excluded from the analyses because of regularize task accuracy lower than 65%. In contrast, every adult who participated in the behavioral study met the accuracy criterion for the regularize task (the worst adult achieved 84% accuracy overall).

At first pass, then, the age group difference in regularize task accuracy may seem an important caveat, especially considering that small differences in error rate, on the order of < 10%, may precipitate relatively large differences in reaction time (Pachella 1974). Even though I cannot ameliorate this potential confound and/or measure its contribution at the level of regularize task performance writ large, it is reassuring that on a finer scale (i.e., considering separately computed main effects and interactions), reaction time patterns were generally paralleled by corresponding accuracy patterns. In other words, as reported in Chapter 3, faster reaction times tracked with higher accuracy for the individual interactions of interest (i.e., task x

lexicality x age group). Also, the faster but less accurate regularize task performance for the

children differed qualitatively from the pattern observed in the read task, in which children were,

as is typically found (e.g., Kail 1991), both slower *and* less accurate than adults

Future Directions

A well-formed body of work achieves its shape by dint of separation of the wheat from

the chaff (or, as Steve Petersen would exhort, "You've got the block of marble; show me

David[14]!"). Sometimes there is substance in that chaff, and in the case of this thesis work,

several lines of inquiry remain untapped. Consideration of the following additional analyses may

prove useful for substantiating and/or refining the interpretation of the data presented in these

chapters.

Reaction time distributional analyses

One approach that may yield additional evidence for or against the 2-mechanism

account is a reconsideration of the reaction time data that accounts for not only the condition

means but also the shapes of the reaction time distributions. In particular, information exists

beyond the mean of a reaction time distribution that can be informative as to what kinds of

processing underlie the task being performed (e.g., Ratcliff 1979; Balota et al. 2008). Meaningful

features of reaction time distributions can be illuminated by fitting the observed data to a

mathematical function (e.g., ex-Gaussian, Gamma, Weibull), and/or by plotting the data using

techniques such as quantile or Vincentile plots (e.g., Andrews and Heathcote 2001; Balota and

Yap 2011). While a full exploration of these complex analytic approaches is outside of the scope

of this work, I will explain briefly how one approach, in particular, warrants consideration for

follow-up investigation.

---

[14] Not Balota.

An ex-Gaussian distribution (composed of a Gaussian convolved with an exponential form) fits many empirically-derived reaction time distributions quite well (Ratcliff and Murdock 1976). Characterized by 3 parameters (mu and sigma for the mean and standard deviation, respectively, of the Gaussian, plus tau to account for the mean and standard deviation of the exponential), ex-Gaussian distributions are specifically useful for capturing the existence of a small number of trials with very slow reaction time, i.e., a tail at the end of the distribution (Balota and Yap 2011). In the present dataset, and consistent with much of the literature, the very slowest trials were removed prior to the calculation of the mean, in order not to bias the calculation of the subject's "average" performance. However, performance may be particularly slow for a given trial for reasons other than a temporary lapse of attention, for example, and especially because in the present dataset I was interested in comparing conditions and groups, the trials in tail of the reaction time distributions may indeed prove critical for more precisely evaluating the individuals' and groups' chronometric data.

One could imagine modeling ex-Gaussian distributions for each stimulus type for each task and each age group separately. Specifically with regard to the proposed response check, one could then directly test the hypothesis that the high frequency regular word trials in the regularize task involve some relatively greater number of trials in which responses are particularly slow (e.g., relative to high frequency irregular words trials and low frequency regular word trials). Such a condition effect would manifest as a greater value for tau (the exponential parameter of the distribution) in the high frequency regular word condition, as compared to all other conditions. Given the fact that the correlation analyses in Chapter 3 suggested that children do not show the same frequency effect reversal for regular words for the regularize task as did the adults (see Figure 3.5), one might furthermore predict that observation of a relatively larger tau parameter for high frequency regular words would be more pronounced in the adults than in the children.

The pseudoword task effect would also be interesting to assess using ex-Gaussian analyses. Because it is particularly intriguing that both adults and children, but especially adults, were slower to regularize than to read pseudowords, a distinction between a shift in the means (mu) of the distributions and/or an increase in the tails (tau) may clarify whether the regularize task slowdown derives from 2 distinct sources (possibly manifest as a shift in both mu and tau) or on source (a shift in one only component). Obviously, of particular interest would be whether those relative contributions are different in children and adults. Certainly, the entire set of reaction time analyses could be recomputed using an ex-Gaussian or some other model of the distribution, and such an approach would likely uncover additional features of the data beyond the predictions hazarded here.

Subsidiary imaging analyses: trial sorting by reaction time

Related to the reaction time distributional analyses are subsidiary imaging analyses that may be able to furnish additional evidence in support of the 2-mechanism model. Specifically, because the imaging study used an event-related design, it should be possible to sort the regularize task trials according to response times, then compare those trials with fast reaction times vs. those that with slow reaction times. One prediction that could be informed by reaction time trial sorting involves identifying all trials of a single stimulus type and comparing those with the slowest reaction times against those with the fastest reaction times. Taking the example of high frequency, regular words, one could ask whether the slowest trials show greater activity in the left Frontal-parietal (LFP) regions than the fastest trials, as would be predicted by the response check step in the 2-mechanism model. Relatedly, a discovery that the differential activity for slow vs. fast trials in the LFP regions is greater than the corresponding differential activity in Cingulo-opercular /Dorsal attention + (CO/DA+) regions would provide further support for the characterization of the LFP regions as playing a later, more decision-like role in the

attentional control of reading (as compared to an earlier, less involved role for the CO/DA+ regions in pathway control).

An obvious objection to the proposition of trial sorting by reaction time is that a firm understanding of the effect of reaction time on the BOLD signal is yet to emerge from the neuroimaging literature. It has been suggested that reaction time exerts regionally specific effects on BOLD activity, in addition to being manifest as either a delay and/or an increase in the amplitude of the hemodynamic response (Yarkoni et al. 2009). Some authors have gone so far as to suggest that the contribution of time on task to the BOLD signal is so great in particular regions that the activity in that region cannot rightly be attributed to cognitive processes of interest (e.g., see Grinband et al. 2011 for an argument of this sort regarding dorsal medial frontal cortex). There is also evidence that the relationship between the BOLD signal and reaction time cannot reliably be assumed to be identical between groups, for example children and adults (Carp et al. 2012). Nonetheless, many investigators rely on the first order approximation that reaction time effects on the BOLD signal are linear (e.g., Weissman et al. 2006), an assumption that underlies the occasional methodological practice of regressing reaction time out of data prior to analysis. One potential pitfall of linear RT regression, however, is the fact that longer reaction times could be associated with *more* processing in a region, *different* processing in a region or *both*, a distinction that seems fundamental.

My best guess is that both latency to respond and time on task do contribute in complex ways to the "processing" BOLD signal of interest, in some cases perhaps causing type I or type II errors with regard to particular analyses. However, until a clearer consensus is reached in terms of what reaction time does and how the potential complexity of its effects can be separately identified, I am generally inclined to leave reaction time in the data, at least for the "primary" analyses of a study. That being said, for the sake of completeness, the results of a subsidiary analysis of the imaging data in which reaction time was regressed out are

summarized briefly in Supplemental material A.3a; the take-home is that our essential findings were unchanged by RT regression. Relatedly, I would consider it worthwhile to explore the possibility of sorting trials based on reaction time to assess the predictions described, despite the caveats.

Switch (lexicality) vs. stay trials: examining local context effects

Another approach to examining the imaging data that may yield additional insights involves querying for BOLD signal differences between trials of a particular stimulus type that are presented consecutively vs. those that are presented interleaved with other stimulus types. Trials from the imaging study could thus be sorted into bins of "switch" (e.g., a word followed by a pseudoword) and "stay" (e.g., a word followed by another word), and then compared against one another. The logic here would be reminiscent somewhat of the list composition studies mentioned earlier (e.g., Monsell et al. 1992; Jared 1997; Lupker et al. 1997), whereby effect sizes dependent on stimulus blocking (as compared to stimulus mixing) may reflect strategic deployments of attentional control that are only efficient or possible at a time scale longer than one individual trial. Although I am not proposing implementing pure vs. mixed blocks, the idea in common between that approach and a switch vs. stay comparison is that context effects may be influential on performance.

Because of the magnitude and reliability of the lexicality effect in both the behavioral and the imaging data, one obvious approach would be to label a trial as switch or stay according to whether or not the lexical status of the stimulus changed with regard to the immediately previous item (as in the examples, above). A comparison of switch vs. stay trials may reveal corroborating (or not) evidence of the 2-mechanism account of the attentional control of reading.

For example, the 2-mechanism account may predict greater BOLD signal modulation by item-specific characteristics in the CO/DA+ regions' BOLD signal than the LFP regions', manifest as a greater interaction of trial type (swich vs. stay) with lexicality and time in the

147

CO/DA+ regions than the LFP regions. This prediction would follow from the fact that the nature

of the response checking for any individual stimulus is presumably so dependent on the

particular features of that stimulus that the LFP response should be less affected, relative to the

CO/DA+ readout from pathway control, by the stimulus or stimuli immediately prior to the target

item. That is, if LFP activity indeed reflects a very specific evaluation between the pronunciation

usually associated with a written word and the pronunciation that is soon to be uttered, the

response checking that occurred for the previous item should have little bearing on the need for

checking on the current trial. Activity in CO/DA+ regions, hypothesized in Chapter 2 to read out

the efficacy of the pathway control for a particular combination of stimulus and task, may instead

be predicted to show greater sensitivity to context effects. Although this prediction is tentative, it

is consistent with our conceptualization of trial-related activity in the Cingulo-opercular core

regions as pertaining to a sort of reporting signal, possibly also including information relevant to

performance feedback (Neta et al. *under revision*), and thereby likely more sensitive to

processing that occurred immediately previously.

Comparing task-sustained signals in the regularize task vs. the read task

One inroad to clarifying the nature of the trial-by-trial contribution of the core Cingulo-

opercular regions of the CO/DA+ cluster to read and regularize task performance is to

implement the read vs. regularize comparison using a mixed block/event-related study design

(as reviewed in Petersen and Dubis 2012). Mixed block/event-related designs exploit particular

trial timings and task blockings in order to parse out task-level effects (e.g., task maintenance

signals, reflecting the processing parameters required for task completion) from trial-level

effects (e.g., error signals). Especially given the discussion in Chapter 2 (and alluded to above)

regarding the yet-elusive characterization (Neta et al. *under revision*) of the trial-related BOLD

activity in core Cingulo-opercular regions (defined by Dosenbach et al. 2006), I would be

especially keen on using the mixed block/event-related design to tease apart possible

contributions from the anterior cingulate and bilateral anterior insula regions over different time scales. Of course, a mixed design could obviously be leveraged more broadly to compare the read and regularize tasks, in that differential contributions to the 2 tasks in the sustained signal could be observed in other regions in informative ways.

Exploring semantic processing in the default mode network

An orthogonal issue in the reading and language literature writ large concerns the issue of whether or not, and/or to what extent, regions comprising the default mode network (Raichle et al. 2001; Gusnard et al. 2001; Buckner et al. 2008; Fair et al. 2008) contribute to the processing of semantics. For example, it has been proposed that the consistent observation of activity in the left angular gyrus during reading and language tasks results from that region's role in semantic analysis (Demonet et al. 1992; Binder et al. 2005; Frost et al. 2005; Mechelli et al. 2007). Binder and colleagues (e.g., Binder et al. 2003; Binder et al. 2005) have forwarded a rather strong interpretation of the evidence that the left angular gyrus shows activity related to semantics, in particular suggesting that the region "occupies a position at the top of a processing hierarchy underlying concept retrieval and conceptual integration" and that it is "involved in all aspects of semantic processing" (Binder et al. 2009). In a recent survey of the literature, Seghier gathered evidence supporting a more nuanced and heterogenous functional ascription of the angular gyrus (Seghier 2013). In broad strokes, dorsal portions of the region (possibly encompassing parts of posterior inferior parietal lobule) were suggested to contribute more so to bottom-up types of processing, while ventral portions (probably more overlapping with classic, deactivating default mode network territory) were suggested to participate in top-down types of processing. Per that author's analysis, the ventral portion of the angular gyrus may perform computations related to self-referential processing, computations that, in theory, could provide constraints on semantic analysis (i.e., rather than necessarily being involved in all aspects of semantic processing, as per Binder and colleagues).

In this laboratory, we have been hesitant to define a task-general role for the left angular gyrus, in part because the region tends to show task-induced deactivations and BOLD deactivations have historically provoked a greater degree of interpretive hand-wringing than have, for example, *positive* event-related deflections of the BOLD signal (but see Sestieri et al. 2010 for an exception in adults, or Church et al. 2008 for evidence of positive angular gyrus activity in children). One study from our group showed, using hierarchical clustering, that a left angular gyrus region was unique in having the overlapping properties of a negative dynamic range and sensitivity to lexicality (Church et al. 2011). In that study, the task was similar to our read task (simple speeded naming), but the lexicality manipulation was relatively weak, in that the words were all very low frequency (and subjects were not queried post-scanning to ascertain familiarity with the word stimuli). Relatedly, Seghier and Price looked indirectly for involvement of default mode network regions in semantic processing by examining activity related to a contrast of semantic and perceptual mapping tasks on the one hand (to identify semantic regions) and deactivations relative to fixation on the other (for the default mode networks) (Seghier and Price 2012). Their maps showed overlap in a single region, located in left angular gyrus (MNI -48, -68, +28) close to the region identified in the Church et al. study (2011).

The present dataset was certainly not designed to assess semantic processing directly. However, lexicality was one of the primary stimulus features manipulated in our study, and lexicality is a decent proxy for semantic content. And although neither the read nor the regularize task requires explicit semantic analysis (as is required for category judgment, for example), there is fairly widespread agreement in the literature that semantic properties of words are accessed automatically upon presentation (Neely 1977), even when other tasks are being performed, and even when words are not perceived consciously (Luck et al. 1996). Also, it would be possible to perform a *post hoc* sorting of the words on the basis of imageability or

familiarity ratings, for example, to obtain a finer-grained estimate of the semantic content of the words.

Caveats acknowledged, then, the present imaging dataset could contribute to clarifying the relationship between default mode network activity and semantics. Based on preliminary observations, it seems most likely that these data would identify some level of gradation among default mode regions in terms of sensitivity to lexicality. For example, in one of the iterations of the cluster analysis described in Chapter 2, I noticed 3 sub-regions along the left angular gyrus that showed progressively greater sensitivity to lexicality in the read task. Following up on that observation with more detailed analyses may allow the delineation of patterns, within subsets of default mode network regions, showing task-invariant sensitivity to lexicality (i.e., lexicality effects in both the read and regularize tasks that are in the same direction) vs. task-dependent sensitivity to lexicality (i.e., lexicality effects in the two tasks that are in opposite directions, as was observed in many control-related regions in Chapter 2). The former observation (task-invariance of the lexicality effect) would more strongly suggest a role in semantic processing, especially in the proposed scenario of assuming that the semantic activation that is occurring is automatic and not based on any of the processes required by either task. On the other hand, a finding of task-dependence in the direction of the lexicality effect would be more difficult to attribute cleanly to automatic semantic processing. Hypothetical default mode regions that show lexicality effects in opposite directions in the 2 tasks may play some role in semantic processing, but perhaps in a way that is more sensitive to modulation by attentional control than hypothetical default mode regions showing task-invariant lexicality effects.

Motor planning and execution as another locus of attentional effects

The preparation and implementation of the motor program necessary for word articulation is one component of lexical processing that has yet been addressed in this thesis, despite its possible relevance. The importance of accounting for motor planning/articulatory-

related processes in overt reading tasks was recognized very early in the literature, as far back as Cattell (1886). Relatively more recently, McCann and Besner wrote about articulatory planning as having to do with the accessibility of "condition-action rules" (1987). It seems reasonable to assert that in the present work, the "condition-action rules" used to generate the pronunciations in the regularize task (presumably completely unrehearsed in the case of irregular words) may serve as yet another locus of processing that should be considered (along with pathway control and response checking) as subject to modulation by top-down attentional control signals.

Instead of considering effects of motor planning and execution explicitly, this work has thus far emphasized the *cognitive* processes of pathway control and response checking, even making the point that the process proposed of response checking has been neglected as a potential locus of essential attentional effects. In these final pages of attempting to think broadly about the data and the results, I raise the possibility that the preparation and execution of the articulation may represent yet another neglected stepchild locus of interaction between reading and attentional control, just waiting to be more closely interrogated.

In support of the stepchild notion, in Chapter 2, the left Frontal-parietal (LFP) cluster that was argued to show properties reflecting a role in response checking was also found to sit proximally, in cluster space, to regions in bilateral inferior frontal gyrus *pars opercularis* (IFG*po*), bilateral mouth somatomotor cortex and bilateral cerebellum, regions all strongly suspected to be involved in the motor articulation of speech (e.g., Price 2012; Taylor et al. 2012). The functional similarity of the LFP response check regions and oro-articulation regions that is implied by the dendrogram proximity was argued, in Chapter 2, to support the attribution of the cognitive process of response checking to the LFP regions, in that the IFG, mouth motor and cerebellar regions did not fall into the *same* cluster as the 10 core LFP regions, but instead immediately *adjacent*. Recall that for clustering techniques, the decision of where to draw the

threshold on the dendrogram is arbitrary, although it can certainly be informed (as it was in Chapter 2) by inspection of the relevant pieces. With that arbitrariness in mind, depending on the level at which functional associations are thresholded on the dendrogram, it may be preferable to conceptualize this larger set of regions (including the LFP members and the adjacent motor regions) as one functional unit.

If the critical post-lexical contribution to the regularize task slowdown is less a purely cognitive phenomenon of response checking and more a motor (or combined cognitive-motor) phenomenon of accessing and articulating a response, a behavioral manipulation may provide supporting evidence. One method that has been employed in behavioral studies to tease apart the cognitive aspects of word recognition from the subsequent preparation and execution of the overt response is the delayed naming paradigm (e.g., Forster and Chambers 1973; Theios and Muise 1977; Balota and Chumbley 1985). In delayed naming tasks, participants are presented with identical stimulus types and task instructions as in speeded naming tasks, except that they are asked to withhold their overt responses until a cue appears on the screen, e.g., brackets encircling the target item. Often, delayed naming tasks include several different intervals of time between stimulus presentation and the cue to pronounce the item, typically in the range of 150[15]-2000 ms. The idea is that any effect that is comparably obtained for both normal and delayed naming reflects process(es) that occur following lexical access. Typically, the post-lexical effects identified by delayed naming tasks are assumed to map onto processes required for response preparation and execution (Balota and Chumbley 1985).

In theory, delayed versions of the read and regularize tasks could be implemented both outside and inside of the scanner, although the hypothetical imaging results would undoubtedly be more complicated to interpret. Most simply, if very similar patterns (behavioral and imaging) were obtained in the standard and delayed task comparisons, it would be necessary to re-

---

[15] Estimates of minimum lexical access time range from 100-300 ms (Balota and Chumbley 1985).

consider the proposed 2-mechanism model, acknowledging the possibility that the motor articulation of a response in the regularize task (distinct from the immediately preceding cognitive process of verifying that response) is an important locus of top-down attentional control. Realistically, a delayed regularize task would likely prove quite challenging to implement, not least because of concerns related to getting the delay interval just right, since delays that are too long may allow subjects to perform rehearsal processes that potentially obfuscate the effect that the delay is intended to reveal (Rossmeissl and Theios 1982).

A related and perhaps more tractable approach to interrogating the comparability (and thus susceptibility to attentional modulation) of the articulatory process for the read and regularize tasks would be would be to measure not only the onset latencies of the responses (as was done) but also the response durations. If the regularize task pronunciation duration for a given item, average across subjects, is larger than the corresponding read task pronunciation for that same item, this finding might support the idea that there is at least *some* contribution to the regularize task slowdown that reflects motor preparation and articulation. Of course, this finding would be especially compelling if it were obtained not only for the irregular words (which generate different pronunciations for the 2 tasks), but also for the regular words and pseudowords.

Comparing children and adults using neuroimaging

Clearly, the data presented in this thesis beg the question as to whether age group differences in the task-associated BOLD activity can be leveraged (1) to elaborate the present suggestion of a 2-mechanism model of the attentional control of reading and/or (2) to characterize more precisely the nature of the developmental differences in read and regularize task performance, especially the relative contributions of pathway control and response checking across the age groups. Indeed, analyses of the child imaging data are the most

154

earnestly and urgently planned future analyses of this thesis, despite having been saved for the final point of discussion.

As an aside – and as alluded to in Supplemental material A.3c – the trajectory from imaging data collection to data usability was a nonlinear one for this study, a fact that contributed to the absence of the developmental imaging study from the pages of this thesis. After our early discovery of a presumptive artifact in the BOLD signal, whereby children were showing consistently lower magnitudes than adults (which is unexpected), we attempted a number of potential fixes, including motion censoring and reaction time regression. Eventually, it was discovered that nuisance signal regression, including removal of the global signal, returned the data to a condition that had higher face validity, specifically producing age group differences that showed a much greater variety of patterns (e.g., children > adults; adults> children; similar activations in the 2 groups). A variety of developmental patterns is consistent both with extant literature and work from our own laboratory. From a starting pool of approximately 50 children, 35 participants furnished usable data, and each of the 35 usable datasets from the children has been pre-processed and is now poised for analysis.

Returning to the planned analyses, it would be interesting to interrogate the activity in control-related regions – including Cingulo-opercular, Dorsal attention and Frontal-parietal regions – to see whether an age-related difference in the patterns of recruitment of these regions clarifies the nature of the large task x age group interaction observed behaviorally. For example, one prediction would be that children would show reduced left frontal-parietal (LFP) cluster activity during regularize task performance than the adults, reflecting a lower overall requirement for the response checking step in the less experienced readers. Additionally, children may show smaller lexicality effects than adults in Cingulo-opercular/Dorsal attention + (CODA+) regions for both tasks, reflecting a reduced ability and/or need to exert substantial inhibition of processing along the lexical pathway. Ideally, the neuroimaging results would

155

provide some leverage in determining the *relative* contributions of pathway control and response checking to the development of the attentional control of reading, since the RT patterns in the 2 age groups suggested, albeit tentatively, a contribution from both mechanisms.

A comparison of children and adults performing the read and regularize tasks using fMRI may also furnish important insights into developmental differences in the attentional control of reading in ways that are totally unpredictable from the results of Chapters 2 or 3. For example, in Chapter 3, it was observed that the groups showed indistinguishable reaction times to regularizing pseudowords (there was a 9 ms difference). Even for effects such as this one, in which no *behavioral* difference is observable between groups, it need not follow that the task-associated neural activity be the same (e.g., Wilkinson and Halligan 2004). In fact, I would be quite interested to see whether a developmental imaging study of read vs. regularize would reveal effects of so-called "behavioral phenocopy" (Schlaggar and McCandliss 2007), since such effects are at least immune from the potential criticism that they are driven by reaction time differences alone (Church et al. 2010).

Just as the developmental behavioral observations discussed in Chapter 3 need not constrain the future developmental imaging analyses, neither should the focus on attentional control-related regions in adults, as presented in Chapter 2, imply that the informative developmental differences will be restricted to attentional control regions. In fact, I would be surprised if it were not possible to identify at least a few noteworthy age group differences in presumptive reading-related regions, including left hemisphere regions in the supramarginal gyrus, inferior frontal gyrus *pars opercularis*, angular gyrus and the putative visual wordform area. For example, it was remarked in Chapter 2 that the adults showed little modulation of the left supramarginal gyrus by task or by stimulus, a finding that we highlighted because it carried forward a thread of investigation into this region's function that has been several years in the making in our own laboratory To follow up on Jessica Church's work on the SMG (Church et al.

2008; Church et al. 2011), I would want to determine, using the developmental imaging data, whether children show increased BOLD activity in this region for the regularize task as compared to the read task, even though we did not obtain task or stimulus effects in the region in the adults. Particularly in the context of a hypothetical scenario in which children show relatively less recruitment of top-down attentional control regions for the regularize task as compared to the adults, the observation of sizeable task effects in the children in the SMG would represent one intriguing instance of 2 groups achieving similar (but not identical) behavior through different neural processing.

Concluding remarks

The studies suggested in this final chapter represent more than intellectual discourses (although they are that, too). Particularly given the discussion in Chapter 1 about the public health significance of dyslexia (Peterson and Pennington 2012), the relatively unsatisfactory options for its treatment (Gabrieli 2009), and the suggestion that deficits of attentional control may factor into at least some proportion of cases (Vidyasagar and Pammer 2010), it is plain that a clearer understanding of reading development is needed. Perhaps some of the future studies I proposed could be useful for addressing that need, for example, by generating hypotheses about ways in which the remediation of reading impairment could be effected by targeting attentional control-related behaviors in the struggling child.

Relatedly, reflecting the alleged dogmas of automaticity and phonology, my initial conceptualization of the read vs. regularize task comparison was that the latter would primarily drive up phonological processing. Accordingly, original hypotheses were forwarded to examine the activity in putative phonological processing regions, including regions in left supramarginal gyrus and left inferior frontal gyrus, presuming that those phonological processing regions would show greater activity in the regularize task than the read task in the adults. Concomitant with the

view of the likely essential sites of modulation of neural activity, our initial predictions for the analyses of the developmental read task behavior included the hypothesis that effects that could be localized to processing along the phonological pathway (e.g., length) would be larger in children than adults, while effects localized to the lexical pathway (e.g., lexicality) would be larger in adults than in children. Although the results were not presented in Chapter 3, one of the remarkable findings of the developmental comparison of the read task alone was that there were very few reliable differences in either the main effects or the interactions of the stimulus characteristics (excepting the item-level significance of the lexicality x age group effect, mentioned earlier). In other words, our leverage for revealing ways in which the children were unlike the adults derived primarily from the manipulation of attentional control (i.e., the comparison of the regularize task to the read task). Likewise, our investigation of fMRI BOLD activity in young adults proved to be most fruitful when we examined regions known to be important for top-down attentional control, rather than the left supramarginal gyrus and left inferior frontal gyrus. Had we either restricted the task design to the read task (rather than juxtaposing it to the attentionally-demanding regularize task) or focused the imaging analyses on "reading-related regions," the story told in these pages would be very different, and likely very much less thought-provoking.

Models of lexical processing that invoke 2 steps/stages/mechanisms, such as the one proposed here, can be found throughout the reading literature (e.g., Atkinson and Juola 1973; Balota and Chumbley 1985; Grainger and Jacobs 1996; Balota and Spieler 1999), hinting at the complexity of the cognitive processes required for reading task performance, combined with the difficulty associated with accommodating a multiplicity of observations with a simple account. In fact, the notion of an abstracted read-out of lexical familiarity has even been explicitly incorporated into models (e.g., Grainger and Jacobs 1996), and invoked as one explanation for

the consistent observation of an inhibitory effect of orthographic neighborhood size[16] on

nonword response latencies in lexical decision tasks (Balota et al. 2004). Another example of a

2-stage model is Balota and Chumbley's conceptualization of the lexical decision task as

involving an initial familiarity/meaningfulness assessment (quick, global), followed, when

necessary, by a slower, more analytic evaluation of the stimulus and whether it can be found in

the lexicon (Balota and Chumbley 1984).

Although a parsimonious account of any observation is, in principle, preferable to a more

complex model, one must also strive, as I have done with this thesis, to embrace the complexity

of one's empirical observations. Echoing the then-provocative argument forwarded by Balota

and Chumbley (1984) questioning the assumption that the word frequency effect in lexical

decision is wholly attributable to the "magical moment" of lexical access (Balota 1990), I would

submit that the present data suggest that an attentional process akin to a decision, indeed

having *"little to do with lexical access,"* is relevant to the regularize task as well. Furthermore,

that process, response checking, interacts with a second attentional control mechanism,

pathway control, to achieve the attentional control of reading, likely in a way that is yet immature

in 8-10 year old readers.

---

[16] An orthographic neighborhood is the set of real words that can be created by changing each letter in an item, one at a time (Balota and Chumbley 1985) (e.g. the orthographic neighborhood for "shirt" includes the following: shirk, shire, skirt, shift and short).

APPENDIX A: METHODOLOGICAL CONSIDERATIONS

Supplemental figure A.1 Depiction of stimulus presentation timing for both studies

   A fixed inter-trial-interval of 1000 ms was used in the behavioral study (Chapters 2 and 3), while a pseudo-randomly jittered inter-trial-interval was used in the imaging study (Chapter 2 only). The inter-trial range of 4,450 – 10,050 ms in the imaging study reflects a fixation period including 1-3 MR frames plus, during the frame in which the stimulus appeared, the 1350 ms prior to and 100 ms following stimulus appearance.
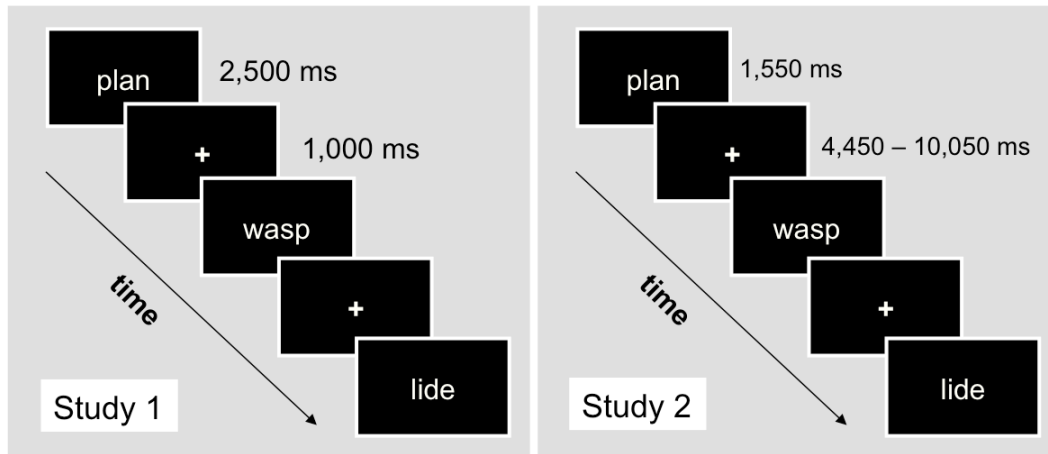
Supplemental figure A.2 Components of one MR frame in imaging study

   Structure of a single, 3 s MR frame in the imaging study (Chapter 2), including the placement of the scanning gap relative to stimulus presentation. The stimulus appeared 1350 ms into the frame, and remained on the screen for 1550 ms. Subjects could respond to a stimulus after its disappearance, provided the next stimulus had not yet appeared.
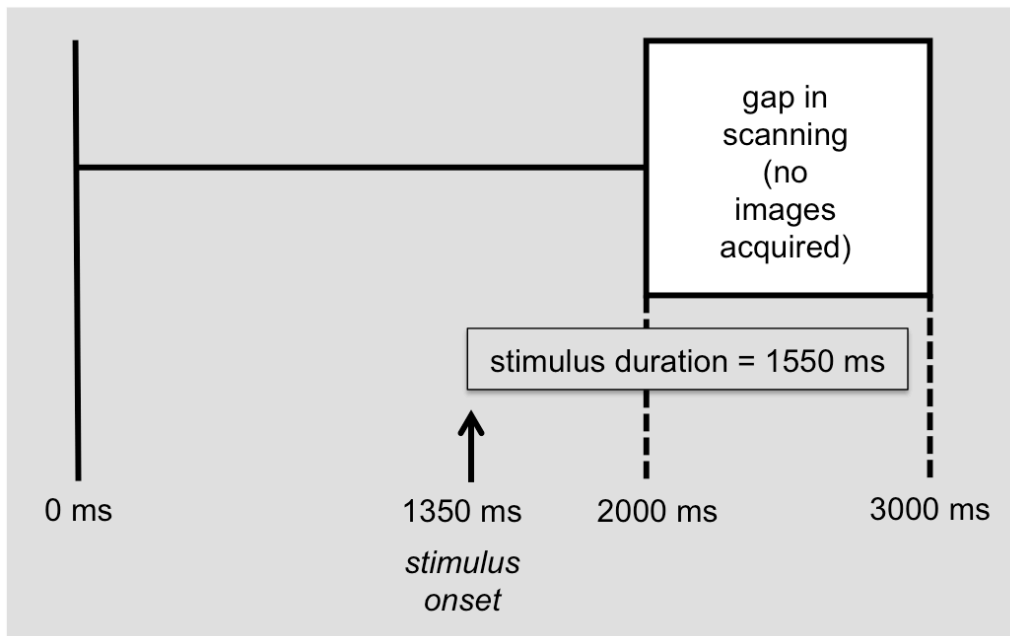
Supplemental material A3: The CO/DA+ vs. LFP segregation in cluster space was generally independent of particular analysis choices.

A3a. Reaction time regression

To address the possible concern that the observed effects were entirely some sort of artifact of response latency and/or time on task, a separate set of GLMs, similar to those described above, was generated for each subject. These GLMs not only included separate terms for errors, stimulus-type, and task, but also coded the reaction time for each individual trial. Reaction time was then treated as a continuous regressor in a supplemental analysis. Specifically, timecourses and statistics were re-extracted, using the regressed GLMs, for the set of regions obtained in the primary analyses. The hierarchical clustering of the original region set was then recomputed.

In the recomputed dendrogram, the exact placement of the 156 regions changed slightly. In fact, the 10 LFP regions remained exclusively and entirely clustered together. While the micro-local relationships between the 20 CO/DA+ were largely intact, the larger cluster broke apart somewhat. However, all 20 regions remained connected at a similarity threshold of $(1-r) = 0.31$, and importantly, they remained segregated (at a distance of $(1-r) = 0.62$) from the LFP region set.

A3b. Motion censoring

The primary imaging analyses in this study were conducted using data meeting a gross movement cutoff of < 1.00 RMS, but without performing any censoring of potential high-movement MR frames (Siegel et al. *in press*). To verify that the primary results were, nonetheless, uncontaminated by movement, timecourses and statistics were re-computed for the set of regions obtained in the primary analyses, using GLMs created from data censored for motion at a framewise displacement (FD) threshold of 0.9 mm (Siegel et al. *in press*).

Motion censoring removed very few frames (on average, 7.5 frames per subject; range 0-46 frames), resulting in qualitative (timecourse patterns) and quantitative (statistical) results comparable to what was obtained using the uncensored data. For example, of 4 statistical effects queried (e.g., task x lexicality x timecourse, etc.) for each of 156 regions (a total of 628 effects), only 10 effects crossed the p = .05 threshold, half of which went from significant to not significant, and half of which went the opposite way.

Partly because frame censoring did so little to change the data, and partly motivated by other factors, we addressed the possibility that motion contributes *some* amount of noise to the measured BOLD signal with an alternate approach, one involving an additional preprocessing step as described in the Methods (Power et al. *in preparation*; see also the next section)

A3c. Preprocessing stream

The Methods describe an additional preprocessing step applied to the data that included the regression of signals from white matter, CSF and whole-brain. Here, we provide additional information about this preprocessing step: the motivation for including this step; a more detailed description of the methodology; and the fact that similar results were obtained when the data was processed in a more traditional manner.

Conceptual and practical motivations

The imaging data reported in Chapter 2 was collected as part of a larger study that included 8-10 year old children. As discussed in the Conclusions, an investigation on the imaging developmental effects is planned as the most proximal next step. Preliminary analyses of the imaging data, however, revealed a signal artifact, presumably due to participant motion, that seemed to affect the children (i.e., higher-movement group) moreso than the adults.

One line of ongoing research in our lab is focused on developing techniques to eradicate the contaminating effects of subject movement on BOLD signal. Recently, we have observed that BOLD signals extracted from voxels within white matter (WM), cerebrospinal fluid (CSF)

and whole-brain (WB) compartments show activity correlated with subtle subject movements (movements that are not captured by grosser estimates) (Power et al. *in preparation*). Furthermore, the application of those WM, CSF and WB signals as "nuisance" regressors during preprocessing seems to increase the quality of both rest- and task-related BOLD data considerably.

Indeed, regressing WM, CSF and WB signal as part of the full preprocessing stream, as described in the present report, ameliorated the age artifact. In order to integrate *this* report with the future developmental report, the primary analyses here were computed using the more involved preprocessing stream, per the Methods. Importantly, the findings of the current study are plausible at several levels (e.g., timecourses have biological shapes, statistical images show believable anatomy, clustering teases apart previously distinguished attentional control systems).

Additional methodological details

To calculate the whole-brain signal, a group-average whole-brain mask was applied. For both the ventricular and the white matter signals, seeds were defined using FreeSurfer segmentations (http://surfer.nmr.mgh.harvard.edu/) of individual subjects' anatomy and then eroded to help ensure that signal obtained from the ventricular and white matter seeds was derived only from the compartment of interest. Four-fold erosions were preferred, but if no voxels remained in the eroded masks, lesser erosions were progressively considered until masks with qualifying voxels were obtained (white matter typically used 4x erosion, ventricular typically 2x or 1x).

Similar results were obtained with alternate preprocessing

Corroborating the face validity of the preprocessing choice, several *post hoc* analyses demonstrated that the fundamental results did not change when the clustering and statistical testing of the 156 identified regions are computed using data that did not undergo the final

preprocessing step (i.e., "un-regressed" data). In the clustering dendrogram, the LFP region

sets stayed entirely and exclusively together at a similarity metric of (1-r) = 0.0625, while the

vast majority (19/20) of the CO/DA+ regions at that same threshold fell into 2 clusters, each of

which gained a handful of new members but nonetheless remained segregated from the LFP

regions.

The data preprocessed in the 2 different ways can even be compared at an *earlier* stage

of analysis, when regions are first derived from voxelwise ANOVAs. Indeed, as seen by

comparing Supplemental figure A.3 with main Figure 2.7, there is much similarity between the

statistical images for the lexicality x task x timecourse interaction obtained using the un-

regressed and regressed data.

Supplemental figure A.3 Statistical image of the voxelwise rmANOVA for the effect of task by lexicality by timecourse, without nuisance regression

The data used to derive this image did not undergo the final preprocessing step described in the Methods. Compare to Figure 2.7, preprocessed with nuisance regression. Color scale corresponds to a *z*-score range of 3.5 to 6.0. Volumetric data was displayed on an inflated cortical surface using Caret software (http://www.nitrc.org/projects/caret/; Van Essen et al. 2001); 3 transverse slices in volumetric space are also shown.
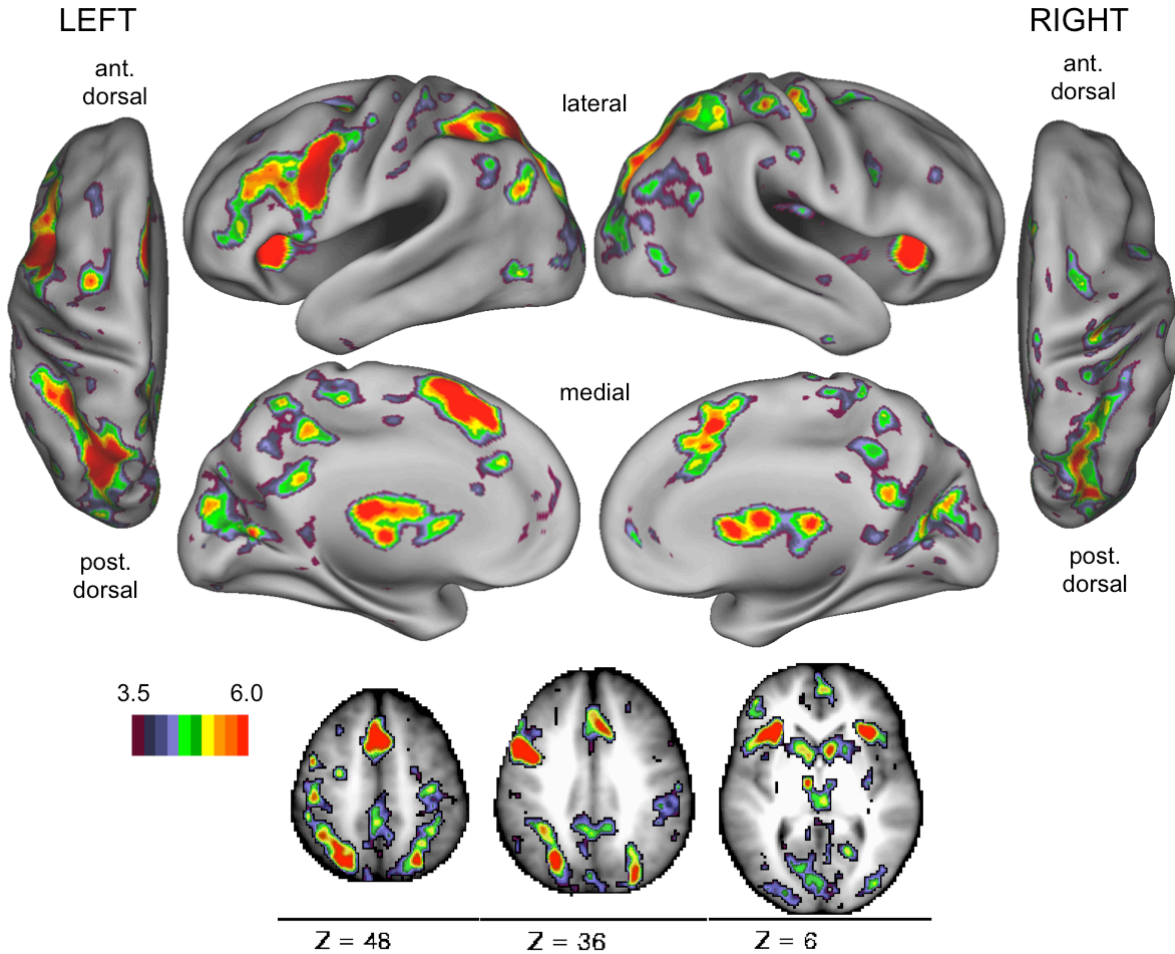
Supplemental material A.4: 156 regions were selected via peak-finding, timecourse extraction, pruning for non-biological timecourses and removal of presumptive duplicates

The automated peak-finding algorithm identified 72 regions of interest (ROIs) in the lexicality x task x timecourse image, and 144 regions in the task x timecourse image, for a total of 216 possible ROIs. Of these ROIs, 7 with non-biological timecourses were removed the lexicality x task x timecourse image, and 32 were removed from the task x timecourse image, leaving 177 candidate ROIs. (A biologically plausible timecourse is one that changes over time (here, 8 MR frames, or 24 s) relatively smoothly, i.e., without vacillating up and down from one timepoint to the next. Many HRF shapes thus qualify as biologically plausible, including a range of positive changes from baseline (activations) and negative changes from baseline (deactivations)).

First-pass UPGMA clustering of the set of 177 ROIs was used to identify rapidly the probable duplicates in the region pool. First, 29 pairs of regions connected by a similarity metric of (1-r) < 0.02 were identified in the dendrogram. Of those 29 pairs, 21 were located within 10 mm Euclidean distance of one another, suggesting that they were likely the same functional region. For each of these 21 pairs, the region with the smaller z score from the parent ANOVA was removed from the pool. This purging resulted in the elimination of 11 regions from the lexicality x task x timecourse ANOVA, and 10 regions from the task x timecourse ANOVA, leaving 156 ROIs. (For the 8 region pairs that were identified as being very close in cluster space ((1-r) < 0.02), but greater than 10 mm Euclidean distance from one another, regions were assumed to be discrete functional entities, thus both members were retained.) The final ROI set therefore included 156 regions, 54 of which derived originally from the lexicality x task x timecourse image, and 103 of which derived from the task x timecourse image.

APPENDIX B: GENERAL STIMULUS INFORMATION

Supplemental material B.1: Complete list of stimuli

B.1a Short stimuli from both the behavioral and imaging studies

Stimuli are divided into 4 word categories and a pseudoword category.
Most items (n=272) were shared across the 2 studies (plain black text).
Items in blue, marked with $, were unique to the behavioral study.
Items in red, marked with #, were unique to the imaging study.

IRREGULAR LOW FREQ	IRREGULAR HIGH FREQ	REGULAR HIGH FREQ	REGULAR LOW FREQ	PSEUDOWORDS	
ache	been	ball $	ace $	afe $	lufts
aisle	blood	best #	batch	arsh	mamp
bury #	blow	black #	beet	aze	marn #
bush $	both $	boy	bleed	bape	mence
calf #	bowl	brick $	braid $	beeg	mip
canoe	bread	camp	cape	bick	mube
chalk	break	check	chant #	bilth	nelk
chef	broad	class #	cheat	blize	nisp
choir	climb	clear $	choke #	boin	nize
chute	come	cloud $	chunk $	breel	nold
comb #	could	coast $	clue	brole	norch
corps	death	damp	craft #	brop	noud $
cough	does	dark	crisp	cag	nug $
crow	done #	day	cube	chen	oabs $
deaf	earth #	duck	drain #	chilf	orp
dealt	flood	face #	duke	chilp #	pice
depot	four $	fence	fate	choy	pide
dose $	front	first #	fetch	clace	pimes
dumb $	give	five $	fig	clox	pirg
echo #	gone	fort	flock #	coaf	pleck
glove #	great #	frog	freak #	com $	plent
gourd $	group	glad $	fudge $	cret	pline
gross #	grow $	grain	gaze	cron	plub
hood	guard	green	gene $	crote	pluft
host $	half	gun $	heel	delk	prale #
hymn	have #	hide	hike	dilp	preef #
limb	heard	hill	itch	doak $	prelk
monk	love	lost	jaw $	drace	prine #
mow	mild $	lunch $	lane	drin	prip

167

ninth	month	**mile $**	lick	**dulp #**	prug
pear	move	more	**loaf $**	elch	purf
pier	none	**name #**	mask	elp	rame
pint	**ocean #**	next	math	feld	raste
plaid	once	nice	**maze #**	felm	rine
pour $	**post $**	**pig $**	**moat $**	flark	rup
sew	prove	pink	ounce	**foint #**	scrut
shove	rough	place	pinch	frabe	shelg
skis	said	plant	plug	fregs	shoin
steak	says	play	quack	frink	slone
swamp	shoe	**point #**	**rib #**	fump	snool
swan	**some #**	prize	saint	funch	solp
swap	soul	reach	scold	gire	storp
swear	soup	**risk $**	shine	gleb	strim
sweat	talk	**road $**	sip	**glort #**	thape
sword	touch	salt	**snout #**	goit	tharl
tomb	tough	shed	spice	**gort $**	thome
ton	two	**side #**	**spin $**	grelp	threp
tour	want	**since #**	stab	grine	tish
tow	**wash $**	space	stink	**gup $**	triss
undo #	wear	**start #**	tame	hace	trube
wad $	what	**state #**	tent	hane	unk
wand	whom	**store #**	thaw	heek	vabb
wasp	won	swim	**trash #**	homp	vame
whoa	youth	tank	trot	horch	varge
yacht		**think #**	**wag $**	jace	vilt
		three #	wink	jelf	vink
		time	witch	joil	voast
		told #	yank	jout	vobe
		train		keeb	weeb
		tribe $		kime	welk
		waste		kirst	woint
		wet $		kolm	wrilk
		wide		krent	wurt
				krong	yain
				leck	zates
				lext	zelf
				lide	zene
				lork #	**zokes #**

168

B.1b Long stimuli (used only in the behavioral study)

Stimuli are divided into 4 word categories and a pseudoword category.

IRREG. LOW FREQ	IRREG. HIGH FREQ	REG. HIGH FREQ	REG. LOW FREQ	PSEUDOWORDS		
anchor	against	asleep	actress	addert	halmint	scitter
ballet	already	attack	artist	agrite	harkete	scrooper
barracks	although	before	banner	alpeer	harvice	scrumple
biscuit	answer	beside	blessing	arpull	hergile	seppad
bishop	become	between	cartoon	belprink	hurgiff	shilmer
bouquet	bicycle	birthday	chipmunk	bentil	huster	shunking
butcher	certain	broken	concrete	billad	imbine	siback
captives	climate	center	contest	bondup	impete	Silfer
chorus	clothing	chapter	declare	bortin	indolt	sillank
circuit	country	chicken	dentist	bripples	inglam	stendeen
cocoon	couple	contract	detect	bruggle	jebail	suppew
combed	cousin	corner	expert	burranks	jiboke	theeger
coyote	daughter	demand	fragment	cabbit	jilden	thorbish
cuckoo	double	describe	goddess	candume	jopper	threndle
cupboard	engine	detail	hamster	capile	jortume	threon
cushion	enough	escape	hunter	carnel	kinter	thrimp
debris	favorite	explain	insane	charfing	korple	trabock
doughnut	feathers	explore	insect	cheolark	kunder	tremair
dungeon	garage	failure	instinct	chilping	laject	untreg
fountain	honest	fishing	invite	chimmer	lectine	varden
fragrant	hungry	forest	ketchup	clindle	lidget	venjures
glacier	husband	golden	kitten	creetul	liggue	veopose
gorgeous	instead	himself	lecture	crenship	linnert	versep
graham	island	hundred	lipstick	crimmoad	lorkit	wockage
granite	league	inside	magnet	cunker	marner	wopple
hatred	machine	kitchen	mattress	degrick	millast	wrundle
heaven	method	market	napkin	dittake	mustick	yapple
horizon	mountain	master	pancake	douncing	namsel	yeggar
jaguar	muscle	mistake	perfume	dreelop	nattipe	zancher
jealous	nature	moisture	pitcher	drince	needure	zurtick
leopard	notice	morning	pumpkin	drussle	octure	
lettuce	palace	myself	raccoon	dushole	orshing	
mammoth	people	outline	reflex	eabine	plaront	
marine	pleasant	percent	remark	ebbult	plendank	
meadow	practice	perhaps	remote	embling	plettle	

169

mischief	private	permit	rescue	estert	plortain
pastry	promise	pressure	sandwich	falmen	poaner
pigeon	purchase	provide	selfish	felken	pocent
pirate	purpose	railroad	sixteen	fointed	prandess
porpoise	service	remain	slender	fortain	preefing
portrait	should	report	splendid	frecker	prendid
ravine	stomach	result	summit	frendure	rapoise
recipe	suggest	sentence	tadpole	frootip	recklune
salmon	surface	subject	temper	fulbec	rellinks
sausage	thousand	surprise	torture	glorking	ruddint
senate	through	thirteen	umpire	grammock	runded
tortoise	toward	transfer	witness	graster	santude
weapon	trouble	winter	zigzag	haggit	scayer

Supplemental material B.2 Examples of irregular test items with pronunciations

Stimulus	IPA transcription[a] (typical pronunciation)	Acceptable regularize task pronunciations
aisle	ajl	ɛsol, ɛslə
anchor	æŋkər	æntʃɔr, ækɔr
barracks	bærəks	bæræks, baræks
broad	brɒd	brod, broæd
could	kʊd	kuld, kɒld, kould
engine	ɛndʒən	ɛndʒajn, ɛngajn
glacier	gleʃər	glesiər, glæsiər, glæʃʊr
guard	gɑrd	gward, guard
pastry	pestri	pæstri, pæstraj, pestraj
recipe	rɛsəpi	rɛsajp, risajp
soup	sup	sawp, soəp
tough	təf	tawəf, tawg, toəf, tof
toward	təwɔrd	tawɔrd, taward, towɔrd, toward
weapon	wɛpən	wipən, wipan, wiəpɑn
yacht	jɑt	jætʃt , jakt

Note. [a]IPA transcriptions guided by http://upodn.com/phon.asp.

Supplemental table B.3 Mean values of lexical properties for 5 short stimulus types
(applies to Chapter 3, as well as Study 1 from Chapter 2)

Stimulus type[a]	Number of letters (SD)	Number of syllables (SD)	Number of orthographic neighbors[b] (SD)	Number of phonological neighbors[b] (SD)	Standard frequency index[c] (SD)
High frequency regular words	4.25 (0.64)	1.00 (0.00)	8.15 (4.89)	15.81 (9.35)	59.51 (5.12)
Low frequency regular words	4.25 (0.64)	1.00 (0.00)	7.46 (5.09)	15.52 (9.91)	46.49 (3.87)
High frequency irregular words	4.33 (0.56)	1.00 (0.00)	6.23 (4.01)	15.52 (9.09)	61.50 (5.96)
Low frequency irregular words	4.29 (0.65)	1.04 (0.20)	5.65 (5.03)	15.27 (10.91)	46.55 (3.60)
Pseudowords	4.29 (0.62)	1.00 (0.00)	5.34 (4.48)	---	---

Note. [a]For all word stimulus types, n = 48; for pseudowords, n =126. [b]Orthographic neighbors and phonological neighbors calculated using the English Lexicon Project database (http://elexicon.wustl.edu; Balota et al. 2007). [c]Standard frequency index (SFI) is from Zeno et al. (1995).

Supplemental table B.4 Mean values of lexical properties for 5 long stimulus types
(applies to Chapter 3)

Stimulus type[a]	Number of letters (SD)	Number of syllables (SD)	Number of orthographic neighbors[b] (SD)	Number of phonological neighbors[b] (SD)	Standard frequency index[c] (SD)
High frequency regular words	6.77 (0.78)	2.00 (0.00)	0.90 (1.48)	1.46 (2.63)	59.17 (3.85)
Low frequency regular words	6.79 (0.74)	2.00 (0.00)	0.81 (1.25)	1.25 (2.05)	47.95 (3.68)
High frequency irregular words	6.79 (0.77)	2.00 (0.36)	0.40 (0.68)	2.46 (3.87)	59.99 (4.77)
Low frequency irregular words	6.75 (0.82)	2.02 (0.33)	0.52 (1.09)	1.46 (2.32)	46.94 (3.38)
Pseudowords	6.76 (0.76)	2.01 (0.09)	0.61 (1.11)	---	---

Note. [a]For all word stimulus types, n = 48; for pseudowords, n =126. [b]Orthographic neighbors and phonological neighbors calculated using the English Lexicon Project database (http://elexicon.wustl.edu; Balota et al. 2007). [c]Standard frequency index (SFI) is from Zeno et al. (1995).

Supplemental material B.5a Items used for demonstration and practice in the behavioral study (n = 13)

Demonstration stimulus	IPA transcription[a] (typical pronunciation)	Regularize task pronunciation (as demonstrated)	Stimulus type
done	dən	don	irregular
have	hæv	hev	irregular
luncheon	ləntʃən	ləntʃiən	irregular
monkey	məŋki	mɑŋki	irregular
remind	rimajnd	rimɪnd	irregular
ribbon	rɪbən	rɪban	irregular
show	ʃo	ʃaw	irregular
soot	sʊt	sut	irregular
squad	skwɑd	skwæd	irregular
worm	wərm	wɔrm	irregular
crippen	krɪpən		pseudoword
checkers	tʃɛkərz		regular
wolf	wʊlf		regular

Note. [a]IPA transcriptions guided by http://upodn.com/phon.asp. IPA transcriptions and regularize task pronunciations are identical for pseudowords and regular words.

Supplemental material B.5b Items used for practice in the behavioral study (n = 60)

apple, bake, bandage, blizzard, brumple, bury, calf, captain, chaos, chilp, clock, contact, curtain, diaper, distance, earth, embrace, façade, famine, famous, father, find, folk, glove, great, gross, hartle, hope, jure, justice, kindness, loke, lump, may, measure, merchant, mind, morale, most, necklace, ocean, orange, pine, plate, porcupine, pregnant, ruck, scissors, scrape, some, swab, tongue, tread, treasure, trip, undo, war, word, wounded, young

Supplemental material B.6a Items used for demonstration and practice in the imaging study (n = 14)

Demonstration stimulus	IPA transcription[a] (typical pronunciation)	Regularize task pronunciation (as demonstrated)	Stimulus type
bush	bʊʃ	bəʃ	irregular
child	tʃajld	tʃɪld	irregular
head	hɛd	hid	irregular
mind	majnd	mɪnd	irregular
soot	sʊt	sut	irregular
squad	skwɑd	skwæd	irregular
tread	trɛd	trid	irregular
wad	wad	wæd	irregular
word	wərd	wɔrd	irregular
bort	bɔrt		pseudoword
ruck	rək		pseudoword
zanch	zæntʃ		pseudoword
clock	klak		regular
risk	rɪsk		regular

Note. [a]IPA transcriptions guided by http://upodn.com/phon.asp. IPA transcriptions and regularize task pronunciations are identical for pseudowords and regular words.

Supplemental material B.6b Items used for practice in the imaging study (n = 69)

ace, answer, bake, ballet, below, bishop, both, bulp, cand, captain, chaos, chirt, chunk, clear, debris, does, double, dumb, find, four, glab, gock, gourd, granite, grow, hope, host, instead, into, island, jibe, jure, jurt, loke, may, meadow, mild, moat, most, people, pig, pine, plate, post, pour, private, recipe, road, runds, salmon, shoff, show, shown, spin, swab, than, treg, tribe, trip, war, wash, weapon, were, wolf, woman, worm, would, young, your

Supplemental material B.7 Frequency did not co-vary with length (applies to Chapter 3)

1.) High frequency words *also* had higher SFIs than low frequency words when each length category was considered separately: short: $t(165) = 20.3$; $p < .001$; long: $t(190) = 21.2$; $p < .001$.

2.) Regular and irregular words were *also* matched for SFI when each length-specific frequency category was considered separately: short low: $t(94) = -0.068$; $p = .946$; short high: $t(94) = -1.755$; $p = .082$; long low: $t(94) = 1.400$; $p = .165$; long high: $t(94) = -0.928$; $p = .356$.

Supplemental figure B.8 Words were chosen to sample a frequency continuum

The high and low frequency words represent a median split of a continuous sampling across the Zeno et al. (1995) Standard Frequency Index (SFI) (shown for the behavioral study words only, since the words used in the imaging study overlapped almost entirely with the short words of the behavioral study).
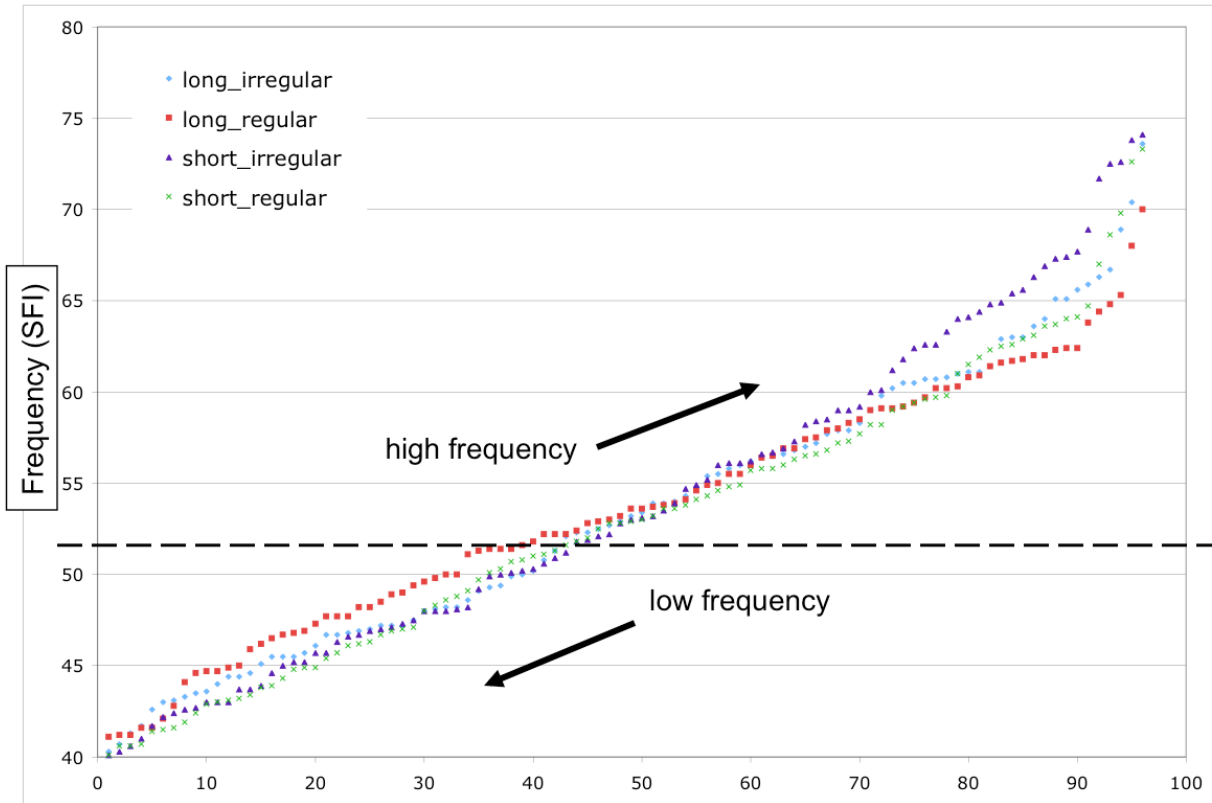
APPENDIX C: SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Supplementary material C.1: A speed advantage for words vs. pseudowords in read is predictive of the lexicality effect reversal in regularize, in an unexpected way.

One alternative to the 2-mechanism account of attentional control interacting with reading processing is a single-mechanism account involving pathway control. A 1-mechanism account in which pathway control sufficiently explains the data might predict that the magnitude of the speed *advantage* for words in the read task would be directly proportional to the magnitude of the RT *slowdown* for words in the regularize task. That is, to the extent that insufficient pathway control *early* in processing entirely accounts for the lexicality effect reversal in regularize task, the strength of the (interfering) lexical route should bear some relationship to the RT benefit afforded to words, compared to pseudowords, in normal reading (see schematic in Supplemental figure C.1). In order to test this particular prediction of a single-mechanism, pathway control hypothesis, we computed a subject-level correlation between the lexicality effect in the read task vs. the lexicality effect in the regularize task in all Study 1 subjects.
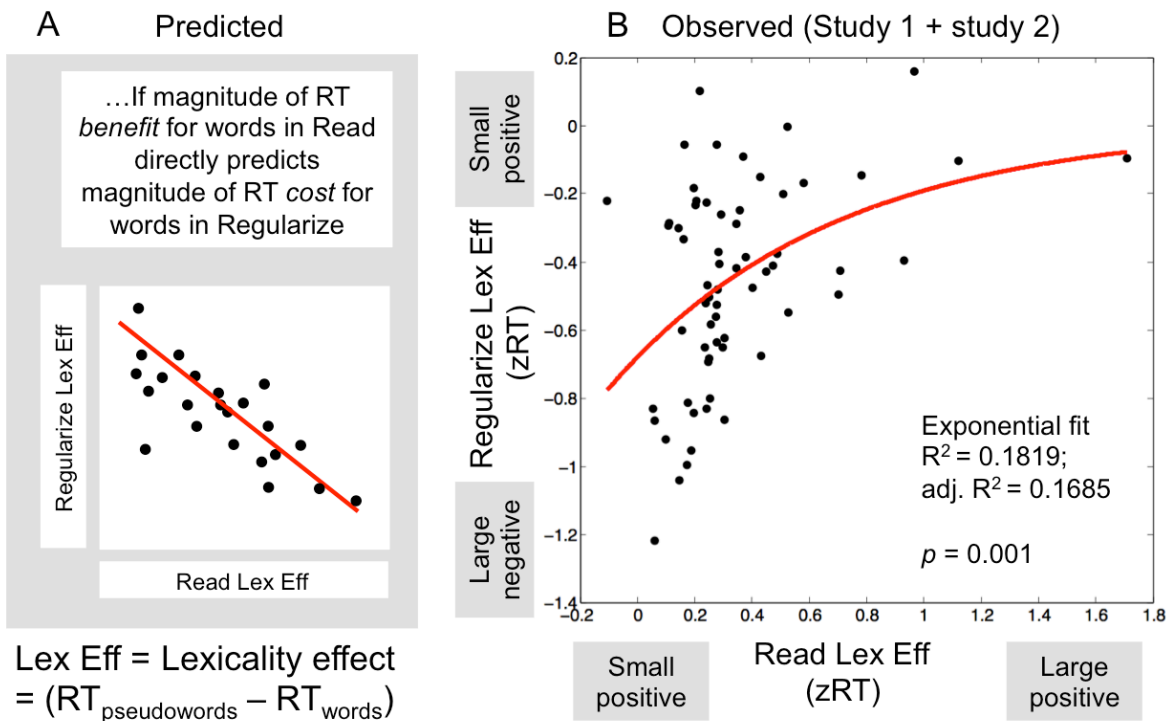
In contrast to a prediction of "equal and opposite" (see main body of the text), the subjects with the *largest* lexicality effect reversals in regularize actually had the *smallest* positive lexicality effects in read (two-tailed p = .013 for the bivariate correlation in Study 1; see Supplemental figure C.1). The relationship was fit by an exponential (R^2 = 0.2125; R^2 adjusted for df = 0.1863), though quadratic, power, Gaussian and even linear fits also explained modest portions of the variance (all R^2 in the range of 0.16 to 0.19). Although, for maximum power, the data points plotted were computed using all the stimuli of both lengths (in the case of Study 1) and both regularities (regular and irregular), similar plots with comparable values of R^2 were obtained when only subsets of the data were considered (e.g., short words only, or excluding irregular words).

The relationship observed in the Study 1 cohort held up in Study 2, as the correlation between small lexicality effects in read and large lexicality reversals in regularize became more reliable in the expanded subject pool (two-tailed p = .001; note that the amount of variance explained by the fit decreased slightly, to R^2 = 0.1819; R^2 adjusted for df = 0.1685; Supplemental figure C.1 shows the combined Study 1 & Study 2 plot).

The relationship between the lexicality effect in the read task vs. the lexicality effect in the regularize task that was observed within individual subjects is challenging to reconcile with an attentional control account that invokes only pathway control. Instead, the incongruity between the size of the lexicality *benefit* for words in read, on the one hand, and the size of the lexicality *cost* for words in regularize, on the other, suggests that there may be downstream contributions of the effect of wordness to the slowing of regularize performance. That is, when skilled readers are asked to compute an unpracticed stimulus-response mapping on a familiar item (i.e., regularize a word), attentional control must be exerted both at the level of achieving appropriate pathway control (accounting for some of the selective slowdown for words) and also some downstream locus (accounting for another portion of the slowdown), quite credibly at the level of response checking.

Supplemental figure C.1 Correlation between lexicality effect in read task vs. lexicality effect in regularize task

Relationship between Lexicality effect in the Read task (x-axis) and Lexicality effect in Regularize task (y-axis) for individual participants. (A) shows schematically the relationship predicted if route biasing (at Mechanism #1) can account entirely for the Task x Lexicality effect, i.e. if the RT *cost* when Regularizing a word is proportional to the RT *benefit* when Reading a word. (B) shows the results actually obtained for individual participants, combined across the behavioral and the imaging study cohorts, as well as the best-fit exponential for the relationship, which explains roughly 18% of the variance. The equation for the fit is $y = Ae^{(Bx)}$, where y = Lexicality Effect in Regularize; x = Lexicality Effect in Read; A = -0.68; B = -1.27.

A — Predicted

…If magnitude of RT *benefit* for words in Read directly predicts magnitude of RT *cost* for words in Regularize

Regularize Lex Eff

Read Lex Eff

Lex Eff = Lexicality effect
$= (RT_{pseudowords} - RT_{words})$

B — Observed (Study 1 + study 2)

Small positive

Regularize Lex Eff (zRT)

Large negative

Exponential fit
$R^2 = 0.1819$;
adj. $R^2 = 0.1685$

$p = 0.001$

Small positive

Read Lex Eff (zRT)

Large positive

Supplemental figure C.2 Surface rendering of the 2 attentional control clusters, with network areal borders

Surface rendering of the CO/DA+ cluster, shown in purple (n = 20 regions), and the LFP cluster, shown in yellow (n = 10 regions). Similar to Figure 2.9, except that network areal borders, adapted from Power et al. (2011), have been added, facilitating visualization of the regions from each of the 2 control clusters within boundaries described using a converging method (rs-fcMRI) and a different set of subjects. Left hemisphere includes borders of the Frontal-parietal control system (yellow), Cingulo-opercular control system (mauve) and Dorsal attention system (green). Right hemisphere includes only Cingulo-opercular and Dorsal attention system borders. Clusters projected using Caret software (Van Essen et al. 2001).
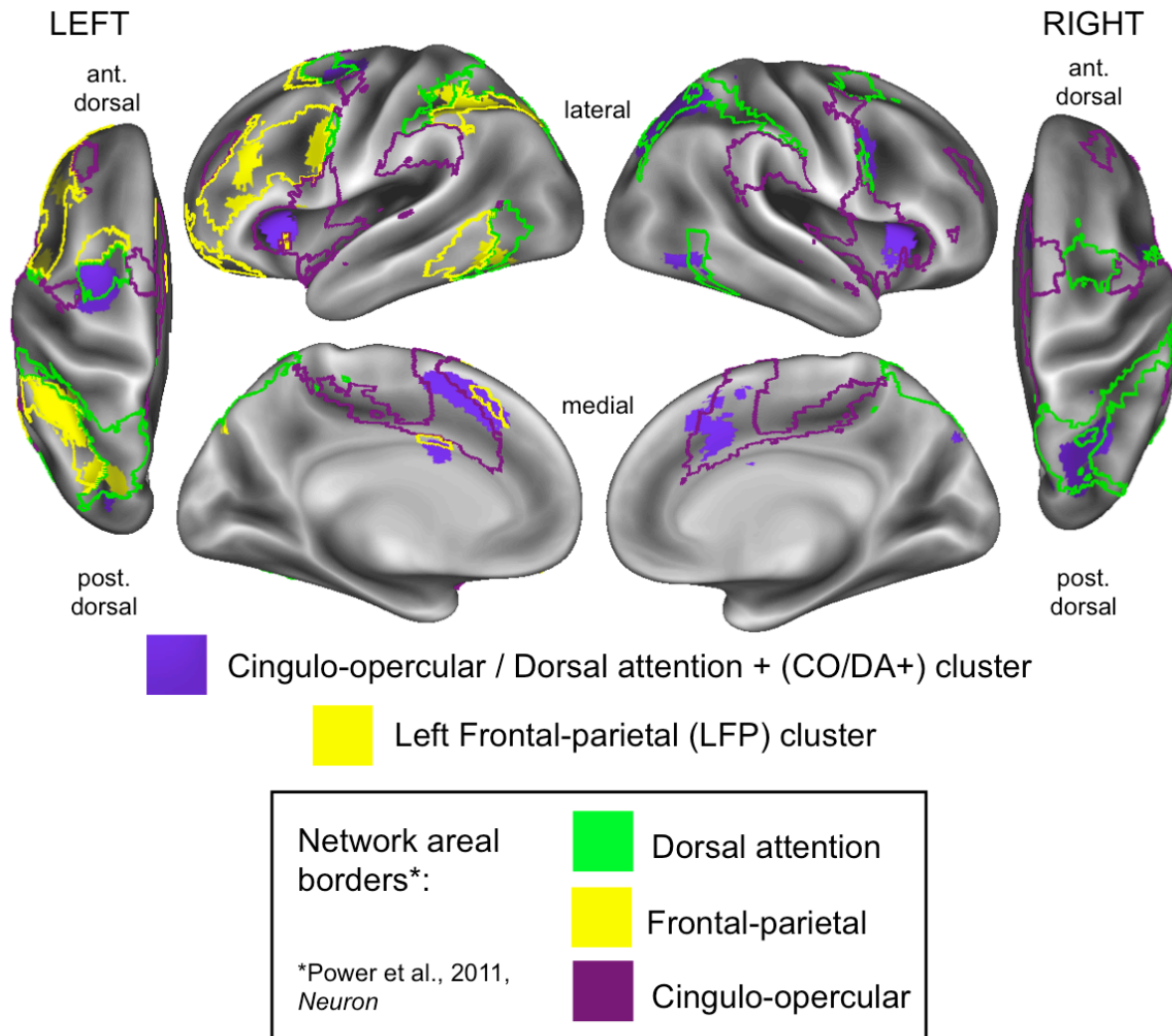
Supplemental figure C.3 Detailed view of dendrogram, focused on LFP cluster

Detailed view of the clustering dendrogram shown in Figure 2.8, here focused on the LFP cluster. Red dashed line indicates clustering threshold of (1-r) = 0.15. X, Y, Z coordinates (MNI) are given for the 10 LFP cluster members as well as their immediate neighbors in cluster space. Note that the 2 closest neighbors are regions in the left and right Inferior frontal gyrus *pars opercularis.* Next most proximal is a cluster of cerebellar regions, followed by a cluster including bilateral mouth somatomotor regions and auditory (STG) regions.
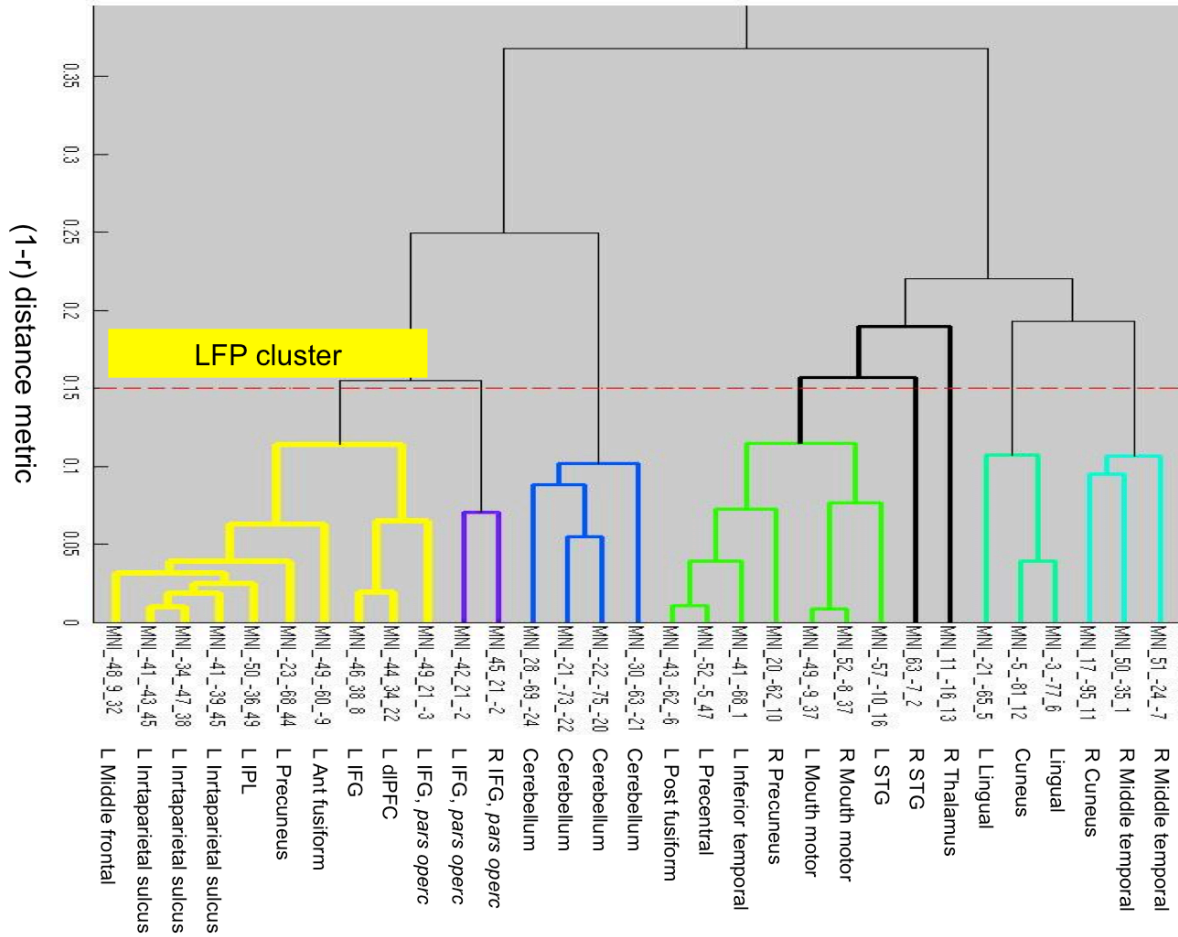
Supplemental figure C.4 Detailed view of dendrogram, focused on CO/DA+ cluster

Detailed view of the clustering dendrogram shown in Figure 2.8, here focused on the CO/DA+ cluster. Red dashed line indicates clustering threshold of (1-r) = 0.15. X, Y, Z coordinates (MNI) are given for the 20 CO/DA+ cluster members as well as their immediate neighbors in cluster space. Note that several subregions of the dorsal anterior cingulate cortex (dACC) are included as distinct members of the cluster. Note also the adjacent lateral visual regions, dorsal attention regions (medial and superior frontal) and caudate regions.
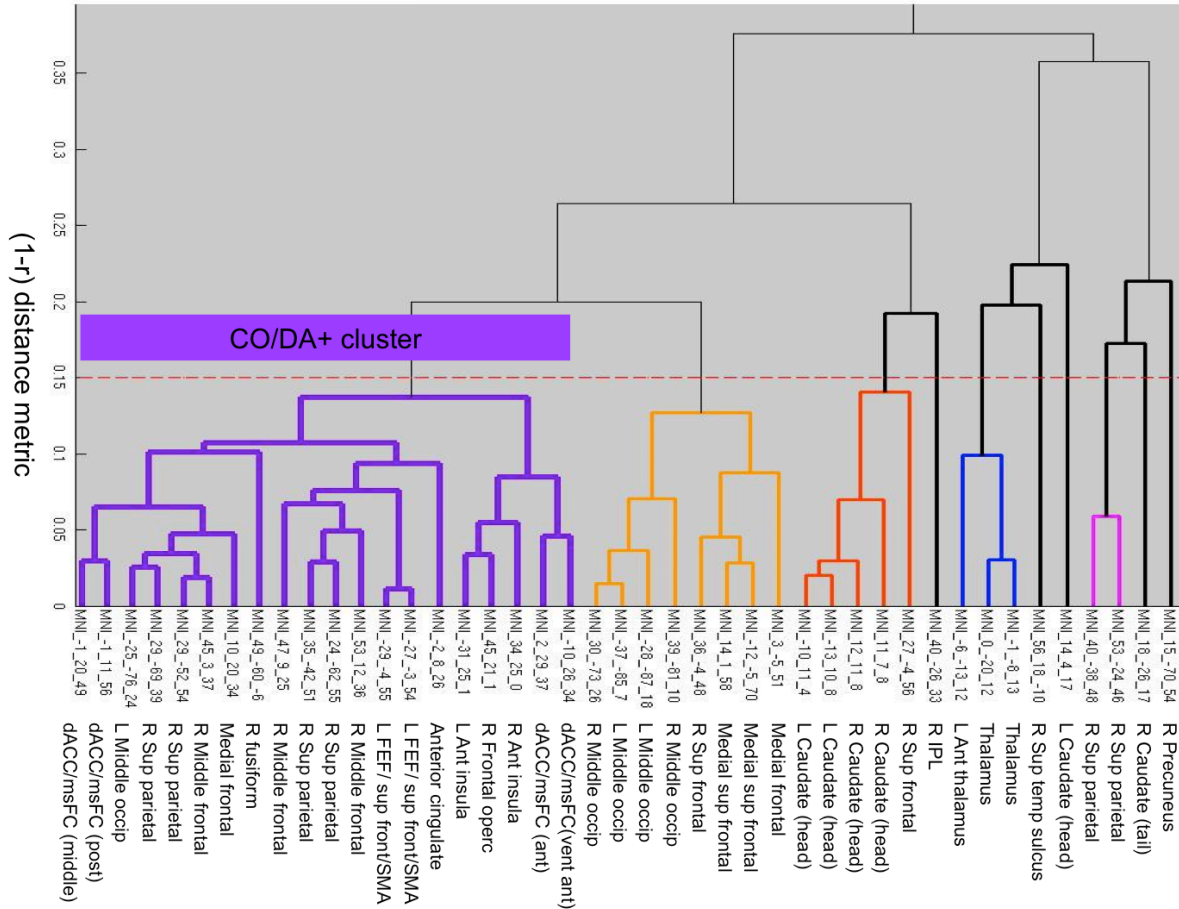
APPENDIX D: SUPPLEMENTARY MATERIAL FOR CHAPTER 3

Supplemental table D.1 Neuropsychological testing results from all tasks except Stroop

Neuropsychological Test	Adult M (SD)	Child M (SD)	Age Group Difference t statistic	Age Group Difference p value[a]
Wechsler Abbreviated Scale of Intelligence (WASI) IQ, percentile rank	97.26 (4.45)	89.92 (12.67)	$t(37.1) =$ -3.047	**0.004**
WASI Vocabulary subtest, ss[b]	16.66 (1.47)	15.10 (2.65)	$t(46.6) =$ -2.87	**0.006**
WASI Matrix Reasoning subtest, ss	13.94 (1.05)	13.16 (1.91)	$t(46.1) =$ -1.99	0.053
Woodcock Johnson III (WJ) Composite reading level, ss[c]	114.81 (8.23)	118.01 (8.77)	$t(61) =$ 1.49	0.140
WJ Letter/Word Identification, ss	118.44 (7.76)	115.94 (8.74)	$t(61) =$ -1.20	0.234
WJ Reading Fluency, ss	121.13 (16.47)	119.90 (14.85)	$t(61) =$ -0.31	0.758
WJ Word Attack, ss	102.41 (9.25)	113.90 (7.48)	$t(61) =$ 5.41	**>.001**

Note. [a]Differences significant at $p < .05$ are bolded. [b]Population range for each WASI subtest is 1-19. [c]Population mean for each WJ subtest is 100.
ss = scaled score.

Supplemental table D.2 Stroop task results

Stroop measure	Adult M (SD)	Child M (SD)	Age Group Difference *t* statistic	Age Group difference *p* value[a]
Facilitation effect (raw RT)	2 ms (54)	74 ms (91)	$t(48.4) = 3.807$	**>0.001**
Interference effect (raw RT)	86 ms (51)	141 ms (114)	$t(41.3) = 2.435$	**0.019**
Facilitation ratio	0.004 (0.077)	0.071 (0.085)	$t(61) = 3.27$	**0.002**
Interference ratio	0.123 (0.074)	0.146 (0.120)	$t(49.8) = 0.92$	0.362
Incongruent zRT – Congruent zRT	-0.055 (0.163)	0.051 (0.163)	$t(61) = 2.60$	**0.012**
Congruent errors	.09 (.30)	.16 (.45)	$t(61) = -0.70$	0.486
Incongruent errors	.53 (.84)	2.48 (2.10)	$t(39.2) = -4.82$	**>0.001**
Neutral errors	.06 (.25)	.65 (.84)	$t(35.0) = -3.72$	**0.001**

Note. [a]Differences significant at p < .05 are bolded.

Supplemental analysis D.2 Examination of Stroop task results

To explore the explanatory adequacy of lexical route interference due to insufficient pathway control for the regularize task slowdown, another point of leverage may be to assess more thoroughly the apparent similarity between the regularize task and the Stroop task. As mentioned in the participant description, the child cohort showed larger Stroop effects (facilitation and interference) than the adults. However, the age group difference in the task effect was in the opposite direction from the group Stroop results, with adults showing greater "interference" from lexical processing than children during the regularize task. Furthermore, individual Stroop measures did not track consistently with measures of read and regularize task reaction times (nor the relationship between the task reaction times[17]), suggesting that the attentional mechanisms underlying the observations of the task x lexicality interaction do not map cleanly onto a Stroop-like interference. Relatedly, Protopapas and colleagues examined the relationship between reading ability and Stroop interference in dyslexic and typically-developing 7th graders (Protopapas et al. 2007). Those authors found that the dyslexic children showed greater Stroop interference than the control children; similarly, within the control group, the poorest readers showed the greatest interference. Our group Stroop results are similar, with the "worse" readers (by absolute reading level, the children) showing larger interference effects (incongruent zRT – congruent zRT; see Table 2) than the "better" readers (the adults). The fact that Stroop task interference is often interpreted as an index of differential pathway strength (e.g., color vs. word naming; see Cohen et al. 1990), coupled with the observation that the regularize task slowdown does not tidily map onto Stroop scores within individuals, reinforces the suggestion, albeit indirectly, that the regularize task slowdown begs for an explanation that invokes something other than *only* lexical route interference due to insufficient pathway control.

---

[17] It is not straightforward to obtain a one-dimensional measure of the task x lexicality interaction; however, we approximated the interaction effect using the following subtraction index: $(pw_{read} - word_{read}) - (pw_{regularize} - word_{regularize})$.

Supplemental material D.3 Calculation of reading levels in both age groups

The children's reading grade level estimates (calculated as a composite of the 3 Woodcock Johnson reading subtests) ranged from 3.1 to 14.9, with an average of 8.0 and a standard deviation of 2.6. On average, the child participants were reading 3.5 grade levels above their actual grade; this sample therefore represents a very skilled group of children. Adult reading grade level estimates ranged from 13.4 to 18.0 (test maximum), with an average of 17.8 and a standard deviation of 0.9; many adult participants were at ceiling in terms of estimated grade level. For statistical comparisons across groups, a reading level *standard score* was computed for each participant, taking into account the participant's actual grade level. The groups did not significantly differ in terms of this grade-adjusted reading level (t(61) = 1.49; p = 0.140). In other words, although the adults had higher absolute reading levels than the children, the two groups did not significantly differ on grade-adjusted reading level. Also, none of the subjects was below the 30th percentile for reading.

Supplemental material D.4 Neuropsychological testing and task performance

correlations of interest

There were a number of notable observations rendered by a consideration of the individual participants' neuropsychological testing scores. Some of those observations are reported here. In total, the following correlations suggest that the best readers (by many measures) are characterized by both *small* positive lexicality effects in the read task and *large* negative lexicality effects in the regularize task. This characterization is consistent with the analysis reported in Supplemental material and figure C.1, whereby the adults with the smallest positive lexicality effects in the read task showed the largest negative lexicality effects in the regularize task.

1) The fastest readers (raw RT for read task) showed the smallest lexicality effects (zRT) for the read task (adults R = .592; $p < .001$; children R = 0.544; $p = .002$).

2) The fastest readers (raw RT for read task) also showed the largest negative lexicality effects (i.e. zRT reversals) in the regularize task; the relationship reached significance in the children and was trend-level in the adults (adults R = .306; $p = .094$; children R = 0.367; $p = .042$).

3) In addition, the fastest regularizers (raw RT for regularize task) showed the largest negative lexicality effects (i.e. zRT reversals) in the regularize task, at least among the adults (adults R = 0.38, $p = .033$; null in children).

4) Higher Letter/Word ID scores (Woodcock Johnson subtest) correlated with both smaller lexicality effects in the read task (adults R = -.586, $p = .001$; children R = -.376; $p = .037$) and larger negative lexicality effects in the regularize task (adults R = -.492, $p = .005$; children R = -.363; $p = .044$).

Supplemental table D.5 Short items only: mean raw reaction times (ms) and accuracy rates

	Read				Regularize			
	Adults		Children		Adults		Children	
Stimulus type	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)
High frequency, regular	672 (119)	1.000 (.000)	822 (122)	0.996 (.012)	1242 (295)	0.989 (.025)	1159 (223)	0.984 (.039)
Low frequency, regular	691 (129)	0.999 (.007)	834 (130)	0.990 (.025)	1209 (295)	0.991 (.028)	1122 (195)	0.977 (.033)
High frequency, irregular	706 (125)	0.995 (.013)	846 (121)	0.953 (.058)	1252 (290)	0.856 (.122)	1204 (241)	0.770 (.108)
Low frequency, irregular	739 (129)	0.970 (.041)	908 (169)	0.847 (.125)	1246 (251)	0.845 (.100)	1193 (262)	0.758 (.095)
Pseudo-words	783 (175)	0.990 (.016)	945 (183)	0.949 (.083)	1081 (277)	0.990 (.022)	1108 (219)	0.968 (.032)

Supplemental table D.6 Long items only: mean raw reaction times (ms) and accuracy rates

	Read				Regularize			
	Adults		Children		Adults		Children	
Stimulus type	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)	RT M (SD)	Acc M (SD)
High frequency, regular	684 (119)	1.000 (.000)	826 (135)	0.993 (.017)	1275 (283)	0.999 (.007)	1120 (208)	0.985 (.027)
Low frequency, regular	692 (122)	1.000 (.000)	838 (135)	0.977 (.044)	1200 (265)	0.996 (.012)	1088 (189)	0.981 (.032)
High frequency, irregular	702 (117)	0.998 (.010)	857 (133)	0.978 (.041)	1296 (268)	0.910 (.102)	1189 (227)	0.773 (.129)
Low frequency, irregular	752 (128)	0.989 (.018)	909 (154)	0.853 (.157)	1288 (284)	0.949 (.066)	1168 (209)	0.839 (.145)
Pseudo-words	902 (227)	0.992 (.014)	1073 (230)	0.904 (.133)	1171 (284)	0.986 (.028)	1163 (211)	0.927 (.091)

Supplemental material D.7 Subject-level correlations between read task and regularize task performance for each subset of stimuli

This supplemental analysis asked whether, within individuals, average reaction time to a particular stimulus type for the read task was meaningfully related to the average reaction time to the same stimulus type for the regularize task. For this analysis, children and adults were considered separately, in order to explore the possibility that there would be age group differences in the nature of the relationship between read and regularize task reaction times for various stimulus types. Of particular interest was the comparison between words and pseudowords, because of the idea that the pseudoword task effect might be attributable largely to an effect of response checking, while the word task effect might result from a combination of effects, in terms of both pathway control and response checking.

The plots show each subject as a dot (adults) or a rectangle (children), and all correct trials for a subject for a particular stimulus type were averaged (collapsing across length). As shown by the top plot, both the children and adults exhibited a linear relationship between pseudoword read task reaction time and pseudoword regularize task reaction time, with very high Pearson correlation coefficients calculated for both age groups (adults R = 0.46, children R = 0.78). In contrast, for all 4 word stimulus types, a clear age group difference was observed, whereby children, but not adults, showed a linear relationship between read task reaction time and regularize task reaction time. The Pearson correlation coefficients as well as the *p* values for the tests of significant linear fit are shown on each plot.

First, it is acknowledged that the finding of a significant relationship between 2 variables in the children coupled with the lack of a corresponding relationship in the adults (as observed for all the word stimuli) does not constitute a direct test of age group difference. However, the correlations seen in the children are compellingly strong for words and pseudowords alike, and given that a linear relationship was observed in the adults for pseudowords, it is worth

speculating about whether the absence of any discernible relationship between read task speed and regularize task speed for adults for words may inform further the results presented in Chapters 2 and 3.

A clear interpretation of the plots shown in Supplemental figure D.7 is made difficult by the fact that within the adult participants, the relationship between the reaction time on the read task and the reaction time on the regularize task not only departs from linearity, it simply falls apart. One point that the plots illustrate well is that for individual children, the response latency for the read task predicts the response latency for the regularize task quite well, for all stimulus types tested. Perhaps a constant increment in the response latency from the read task to the regularize task suggests a major contribution from only 1 mechanism, response checking, in the children, for words and pseudowords. In the case of adults, perhaps the preserved linear relationship observed for the pseudowords likewise indicates the greater contribution to those items' performance from attentional control at the level of response checking rather than pathway control. In that scenario, the nonlinearity of the read vs. regularize task comparison for the word conditions in the adults may reflect a more substantial contribution on those items from pathway control.

It is emphasized that the interpretation just offered to explain the results depicted in Supplemental figure D.7 is highly speculative. One of the approaches that one could take to gain additional leverage on the use of within-subject read vs. regularize task correlations such as those shown is to stratify the word stimuli differently, for example, taking into account length, bigram frequency, orthographic neighborhood size, etc. Some of those variables may illuminate whether certain features of words make them more or less "complicated" in their computation in the regularize task. Similarly, related analyses could be interrogated at the item-, rather than the subject-, level.

Supplemental figure D.7 Subject-level correlations between read task and regularize task reaction time for each subset of stimuli

Each participant is plotted as one marker; red rectangles represent the children and green circles represent the adults. Raw read task reaction times are plotted on the x-axes, while raw regularize task reaction times are plotted on the y-axes. Solid lines indicate a statistically significant linear fit for the correlation, while dotted lines indicate the absence of such a fit.
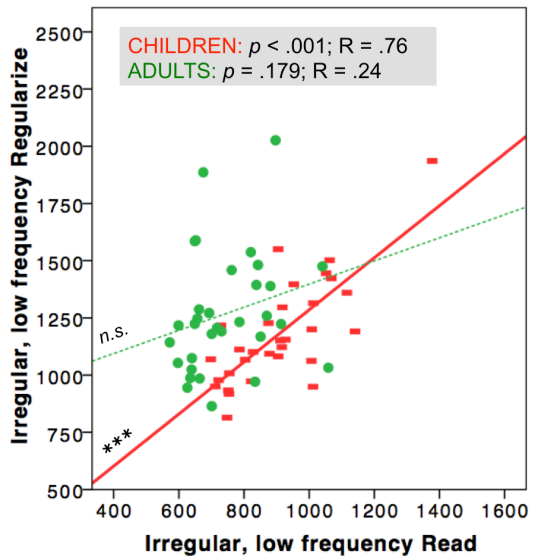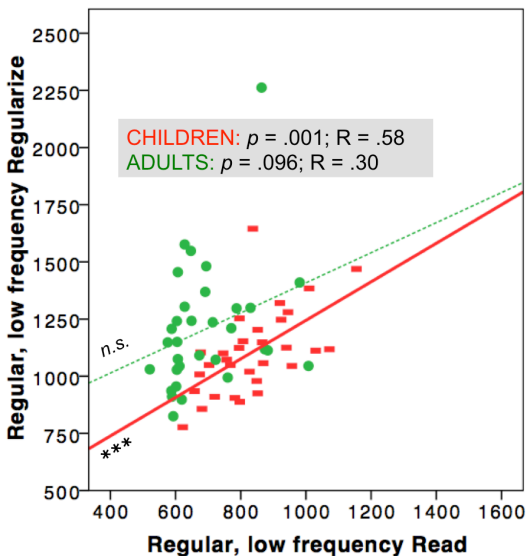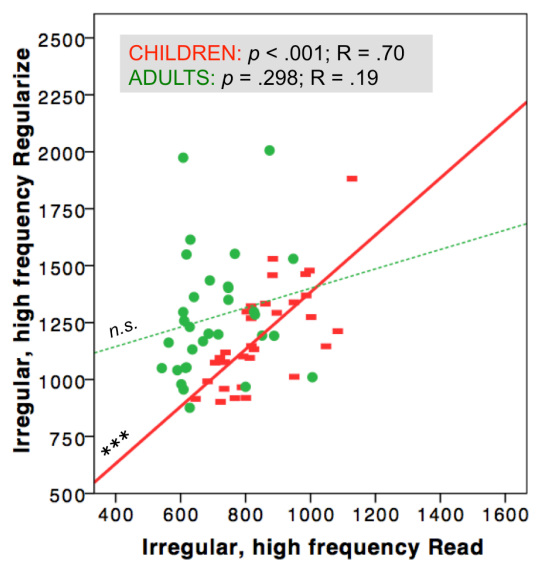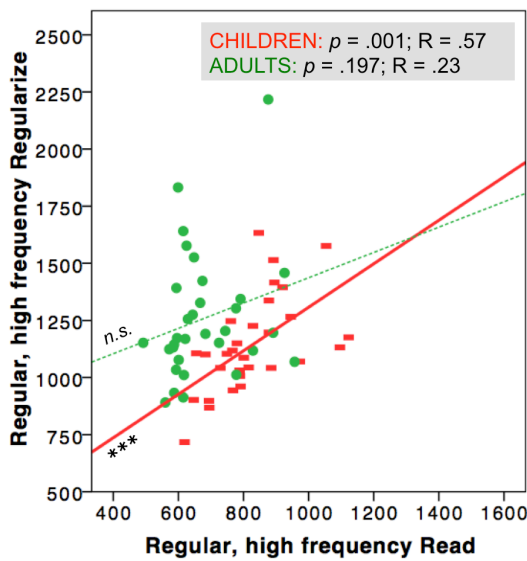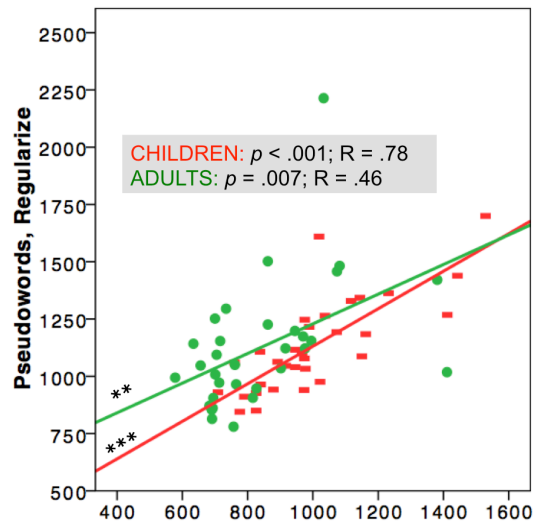
REFERENCES

Aaron PG, Joshi RM, Ayotollah M, Ellsberry A, Henderson J, Lindsey K. 1999. Decoding and sight-word naming: Are they independent components of word recognition skill? Reading and Writing: An Interdisciplinary Journal 11:89-127.

Andrews S. 1989. Frequency and neighborhood effects on lexical access: Activation or search? Journal of Experimental Psychology: Learning, Memory and Cognition 15:802-814.

Andrews S, Heathcote A. 2001. Distinguishing common and task-specific processes in word identification: a matter of some moment? J Exp Psychol Learn Mem Cogn 27:514-544.

Atkinson RC, Juola JF. 1973. Factors influencing speed and accuracy of word recognition. In: Kornblum S, editor. Fourth International Symposium on Attention & Performance New York: New York Academic Press p 583-612.

August GJ, Garfinkel BD. 1990. Comorbidity of ADHD and reading disability among clinic-referred children. J Abnorm Child Psychol 18:29-45.

Backman J, Bruck M, Hebert M, Seidenberg MS. 1984. Acquisition and use of spelling-sound correspondences in reading. J Exp Child Psychol 38:114-133.

Baer J, Kunter M, Sabatini Jeditors. Year Published|. Title|, Conference Name|; Year of Conference Date|; Conference Location| Place Published|:Publisher|. Pages p|.

Balota DA, Chumbley JI. 1984. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. Journal of Experimental Psychology: Human Perception and Performance 10:340-357.

Balota DA, Chumbley JI. 1985. The locus of word-frequency effects in the pronuciation task: Lexical access and/or production? Journal of Memory and Language 24:89-106.

Balota DA. 1990. The role of meaning in word recognition. In: Balota DA, d'Arcais GBF, Rayner K, editors. Comprehension processes in reading Hillsdale, NJ: Erlbaum p 9-32.

Balota DA, Chumbley JI. 1990. Where are the effects of frequency in visual word recognition tasks? Right where we said they were. Comment on Monsell, Doyle, & Haggard (1989). J Exp Psychol Gen 119:231-237.

Balota DA, Cortese MJ, Sergent-Marshall SD, Spieler DH, Yap M. 2004. Visual word recognition of single-syllable words. J Exp Psychol Gen 133:283-316.

Balota DA, Law MB, Zevin JD. 2000. The attentional control of lexical processing pathways: Reversing the word frequency effect. Mem Cognit 28:1081-1089.

Balota DA, Spieler DH. 1999. Word frequency, repetition, and lexicality effects in word recognition tasks: beyond measures of central tendency. J Exp Psychol Gen 128:32-55.

Balota DA, Yap MJ. 2011. Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. Psychol Sci 20:160-166.

Balota DA, Yap MJ, Cortese MI, Watson JM. 2008. Beyond response latency: An RT distributional analysis of semantic priming. Journal of Memory & Language 59:495-523.

Balota DA, Yap MJ, Cortese MJ, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R. 2007. The English lexicon project. Behavioral Research Methods 39:445-459.

Baluch B, Derek B. 1991. Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. Journal of Experimental Psychology: Learning, Memory and Cognition 17:644-652.

Baron J, Strawson C. 1976. Use of orthographic and word-specific knowledge in reading words aloud. J Exp Psychol Hum Percept Perform 2:386-393.

Bench CJ, Frith CD, Grasby PM, Friston KJ, Paulesu E, Frackowiak RS, Dolan RJ. 1993. Investigations of the functional anatomy of attention using the Stroop test. Neuropsychologia 31:907-922.

Binder JR, Desai RH, Graves WW, Conant LL. 2009. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb Cortex 19:2767-2796.

Binder JR, McKiernan KA, Parsons ME, Westbury CF, Possing ET, Kaufman JN, Buchanan L. 2003. Neural correlates of lexical access during visual word recognition. J Cogn Neurosci 15:372-393.

Binder JR, Medler DA, Desai R, Conant LL, Liebenthal E. 2005. Some neurophysiological constraints on models of word naming. Neuroimage 27:677-693.

Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM. 2004. Development of brain mechanisms for processing orthographic and phonologic representations. J Cogn Neurosci 16:1234-1249.

Booth JR, Mehdiratta N, Burman DD, Bitan T. 2008. Developmental increases in effective connectivity to brain regions involved in phonological processing during tasks with orthographic demands. Brain Res 1189:78-89.

Booth JR, Perfetti CA, MacWhinney B. 1999. Quick, automatic, and general activation of orthographic and phonological representations in young readers. Dev Psychol 35:3-19.

Bosse ML, Tainturier MJ, Valdois S. 2007. Developmental dyslexia: the visual attention span deficit hypothesis. Cognition 104:198-230.

Botvinick MM, Cohen JD, Carter CS. 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci 8:539-546.

Boynton GM, Engel SA, Glover GH, Heeger DJ. 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 16:4207-4221.

Brown JW, Braver TS. 2005. Learned Predictions of Error Likelihood in the Anterior Cingulate Cortex. Science 307:1118-1121.

Brown P, Lupker SJ, Colombo L. 1994. Interacting sources of information in word naming: A study of individual differences. J Exp Psychol Hum Percept Perform 20:537-554.

Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, Schlaggar BL. 2005. Developmental changes in human cerebral functional organization for word generation. Cereb Cortex 15:275-290.

Buckner RL, Andrews-Hanna JR, Schacter DL. 2008. The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1-38.

Buckner RL, Raichle ME, Petersen SE. 1995. Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. J Neurophysiol 74:2163-2173.

Bunge SA, Hazeltine E, Scanlon MD, Rosen AC, Gabrieli JD. 2002. Dissociable contributions of prefrontal and parietal cortices to response selection. Neuroimage 17:1562-1571.

Bush G, Luu P, Posner MI. 2000. Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 4:215-222.

Bush G, Whalen PJ, Rosen BR, Jenike MA, McInerney SC, Rauch SL. 1998. The Counting Stroop: An interference task specialized for functional neuroimaging: Validation study with functional MRI. Hum Brain Mapp 6:270-282.

Carp J, Fitzgerald KD, Taylor SF, Weissman DH. 2012. Removing the effect of response time on brain activity reveals developmental differences in conflict processing in the posterior medial prefrontal cortex. Neuroimage 59:853-860.

Carter CS, Mintun M, Cohen JD. 1995. Interference and facilitation effects during selective attention: an H215O PET study of Stroop task performance. Neuroimage 2:264-272.

Casey B. 1997. A Developmental Functional MRI Study of Prefrontal Activation during Performance of a Go-No-Go Task. J Cogn Neurosci.

Casey BJ, Giedd JN, Thomas KM. 2000. Structural and functional brain development and its relation to cognitive development. Biol Psychol 54:241-257.

Castellanos FX, Tannock R. 2002. Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. Nat Rev Neurosci 3:617-628.

Cattell JM. 1886. The time it takes to see and name objects. Mind 11:63-65.
Chateau D, Jared D. 2000. Exposure to print and word recognition processes. Mem Cognit 28:143-153.

Chrysikou EG, Novick JM, Trueswell JC, Thompson-Schill SL. 2011. The other side of cognitive control: Can a lack of cognitive control benefit language and cognition? Top Cogn Sci 3:253-256.

Church JA, Balota DA, Petersen SE, Schlaggar BL. 2011. Manipulation of length and lexicality localizes the functional neuroanatomy of phonological processing in adult readers. J Cogn Neurosci 23:1475-1493.

Church JA, Coalson RS, Lugar HM, Petersen SE, Schlaggar BL. 2008. A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. Cereb Cortex 18:2054-2065.

Church JA, Petersen SE, Schlaggar BL. 2010. The "Task B Problem" and other considerations in developmental functional neuroimaging. Hum Brain Mapp 31:852-862.

Cohen JD, Dunbar K, McClelland JL. 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. Psychol Rev 97:332-361.

Cohen JD, MacWhinney B, Flatt M, Provost J. 1993. PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers 25:257-271.

Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S. 2002. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. Brain 125:1054-1069.

Coltheart M, Curtis B, Atkins P, Haller M. 1993. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. Psychol Rev 100:589-608.

Coltheart M, Rastle K. 1994. Serial processing in reading aloud: Evidence for dual-route models of reading. J Exp Psychol Hum Percept Perform 20:1197-1211.

Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. 2001. DRC: a dual route cascaded model of visual word recognition and reading aloud. Psychol Rev 108:204-256.

Cone NE, Burman DD, Bitan T, Bolger DJ, Booth JR. 2008. Developmental changes in brain regions involved in phonological and orthographic processing during spoken language processing. Neuroimage 41:623-635.

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience 3:201-215.

Crone EA, Bunge SA, van der Molen MW, Ridderinkhof KR. 2006. Switching between tasks and responses: a developmental study. Dev Sci 9:278-287.

Damasio AR, Damasio H. 1983. The anatomic basis of pure alexia. Neurology 33:1573-1583.

Davidson MC, Amso D, Anderson LC, Diamond A. 2006. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. Neuropsychologia 44:2037-2078.

Davies PL, Segalowitz SJ, Gavin WJ. 2004. Development of response-monitoring ERPs in 7- to 25-year-olds. Dev Neuropsychol 25:355-376.

Davis MH, Ford MA, Kherif F, Johnsrude IS. 2011. Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. J Cogn Neurosci 23:3914-3932.

Defior S, Justicia F, Martos FJ. 1996. The influence of lexical and sublexical variables in normal and poor Spanish readers. Reading and Writing: An Interdisciplinary Journal 8:487-497.

Dehaene S, Posner MI, Tucker DM. 1994. Localization of a Neural System for Error-Detection and Compensation. Psychol Sci 5:303-305.

Dejerine J. 1891. Sur un cas de cécité verbale avec agraphie, suivi d'autopsie. Compte Rendus Hebdomadaires Séances et Mémoires de la Société de Biologie 3:197-201.

Dejerine J. 1892. Contribution a l'étude anatomoclinique et clinique des differentes varietes de cecite verbal. Compte Rendus Hebdomadaires Séances et Mémoires de la Société de Biologie 4:61-90.

Demb JB, Boynton GM, Heeger DJ. 1998. Functional magnetic resonance imaging of early visual pathways in dyslexia. J Neurosci 18:6939-6951.

Demonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, Wise R, Rascol A, Frackowiak R. 1992. The anatomy of phonological and semantic processing in normal subjects. Brain:1753-1768.

Dennis I, Newstead SE. 1981. Is phonological recoding under strategic control? Mem Cognit 9:472-477.

Devlin JT, Rushworth MF, Matthews PM. 2005. Category-related activation for written words in the posterior fusiform is task specific. Neuropsychologia 43:69-74.

Dietz NA, Jones KM, Gareau L, Zeffiro TA, Eden GF. 2005. Phonological decoding involves left posterior fusiform gyrus. Hum Brain Mapp 26:81-93.

Doctor EA, Coltheart M. 1980. Children's use of phonological encoding when reading for meaning. Mem Cognit 8:195-209.

Donders FC. 1969. On the speed of mental processes. In: Koster WG, editor. Acta Psychologica 30 Attention and Performance II   North Holland Publishing p 412-431.

Dosenbach NUF, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. Trends Cogn Sci 12:99-105.

Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE. 2007. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104:11073-11078.

Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. 2006. A core system for the implementation of task sets. Neuron 50:799-812.

Eden GF, VanMeter JW, Rumsey JM, Maisog JM, Woods RP, Zeffiro TA. 1996. Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. Nature 382:66-69.

Eden GF, VanMeter JW, Rumsey JM, Zeffiro TA. 1996. The visual deficit theory of developmental dyslexia. Neuroimage 4:S108-117.

Eden GF, Zeffiro TA. 1998. Neural systems affected in developmental dyslexia revealed by functional neuroimaging. Neuron 21:279-282.

Elliott R. 1970. Simple reaction time: effects associated with age, preparatory interval, incentive-shift, and mode of presentation. J Exp Child Psychol 9:86-107.

Fair DA, Cohen AL, Dosenbach NU, Church JA, Miezin FM, Barch DM, Raichle ME, Petersen SE, Schlaggar BL. 2008. The maturing architecture of the brain's default network. Proc Natl Acad Sci USA 105:4028-4032.

Fairweather H, Hutt SJ. 1978. On the rate of gain of information in children. J Exp Child Psychol 26:216-229.

Fan J, McCandliss BD, Sommer T, Raz A, Posner MI. 2002. Testing the efficiency and independence of attentional networks. J Cogn Neurosci 14:340-347.

Faust ME, Balota DA, Ferraro FR, Spieler DH. 1999. Individual differences in information-processing rate and amount: implications for group differences in response latency. Psychol Bull 125:777-799.

Fernandes S, Ventura P, Querido L, Morais J. 2008. Reading and spelling acquisition in European Portuguese: a preliminary study. Reading Writ 21:805-821.

Fiebach CJ, Ricker B, Friederici AD, Jacobs AM. 2007. Inhibition and facilitation in visual word recognition: prefrontal contribution to the orthographic neighborhood size effect. Neuroimage 36:901-911.

Fiez JA. 1997. Phonology, semantics, and the role of the left inferior prefrontal cortex. Hum Brain Mapp 5:79-83.

Fiez JA, Tranel D, Seager-Frerichs D, Damasio H. 2006. Specific reading and phonological processing deficits are associated with damage to the left frontal operculum. Cortex 42:624-643.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. Magn Reson Med 33:636-647.

Forster KI, Chambers SM. 1973. Lexical access and naming time. Journal of Verbal Learning & Verbal Behavior 12:627-635.

Forster KI, Davis ET. 1991. The density constraint on form-priming in the naming task: Interference effects from a masked prime. Journal of Memory and Language 30:1-25.

Fox MD, Zhang D, Snyder AZ, Raichle ME. 2009. The global signal and observed anticorrelated resting state brain networks. J Neurophysiol.

Franceschini S, Gori S, Ruffino M, Pedrolli K, Facoetti A. 2012. A causal link between visual spatial attention and reading acquisition. Curr Biol 22:814-819.

Frederiksen JR, Kroll JF. 1976. Spelling and sound: Approaches to the internal lexicon. J Exp Psychol Hum Percept Perform 2:361-379.

Frost R. 1998. Toward a strong phonological theory of visual word recognition: true issues and false trails. Psychol Bull 123:71-99.

Frost R. 2012. Towards a universal model of reading. Behav Brain Sci 35:263-279.

Frost SJ, Mencl WE, Sandak R, Moore DL, Rueckl JG, Katz L, Fulbright RK, Pugh KR. 2005. A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud. Neuroreport 16:621-624.

Gabrieli JD. 2009. Dyslexia: a new synergy between education and cognitive neuroscience. Science 325:280-283.

Germano E, Gagliano A, Curatolo P. 2010. Comorbidity of ADHD and dyslexia. Dev Neuropsychol 35:475-493.

Giedd JN. 2004. Structural magnetic resonance imaging of the adolescent brain. Ann N Y Acad Sci 1021:77-85.

Glezer LS, Jiang X, Riesenhuber M. 2009. Evidence for highly selective neuronal tuning to whole words in the "visual word form area". Neuron 62:199-204.

Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TF, 3rd, Herman DH, Clasen LS, Toga AW, Rapoport JL, Thompson PM. 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc Natl Acad Sci U S A 101:8174-8179.

Gold BT, Balota DA, Kirchhoff BA, Buckner RL. 2005. Common and Dissociable Activation Patterns Associated with Controlled Semantic and Phonological Processing: Evidence from fMRI Adaptation. Cereb Cortex.

Grainger J, Jacobs AM. 1996. Orthographic processing in visual word recognition: a multiple read-out model. Psychol Rev 103:518-565.

Grainger J, Ziegler JC. 2011. A dual-route approach to orthographic processing. Front Psychol 2:54.

Greenberg D, Ehri LC, Perin D. 1997. Are word-reading processes the same of different in adult literacy students and thire-fifth graders matched for reading ability? J Educ Psychol 89:262-275.

Greenberg D, Ehri LC, Perin D. 2002. Do adult literacy students make the same word-reading and spelling errors as children matched for word-reading age? Scientific Studies of Reading 6:221-243.

Greenlee JD, Oya H, Kawasaki H, Volkov IO, Kaufman OP, Kovach C, Howard MA, Brugge JF. 2004. A functional connection between inferior frontal gyrus and orofacial motor cortex in human. J Neurophysiol 92:1153-1164.

Grinband J, Savitskaya J, Wager TD, Teichert T, Ferrera VP, Hirsch J. 2011. The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. Neuroimage 57:303-311.

Gusnard DA, Raichle ME, Raichle ME. 2001. Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neurosci 2:685-694.

Habib M. 2000. The neurological basis of developmental dyslexia: an overview and working hypothesis. Brain 123 Pt 12:2373-2399.

Hale S. 1990. A global developmental trend in cognitive processing speed. Child Dev 61:653-663.

Hari R, Renvall H. 2001. Impaired processing of rapid stimulus sequences in dyslexia. Trends Cogn Sci 5:525-532.

Hari R, Renvall H, Tanskanen T. 2001. Left minineglect in dyslexic adults. Brain 124:1373-1380.

Hari R, Valta M, Uutela K. 1999. Prolonged attentional dwell time in dyslexic adults. Neurosci Lett 271:202-204.

Harm MW, Seidenberg MS. 1999. Phonology, reading acquisition, and dyslexia: Insights from connectionist models. Psychol Rev 106:491-528.

Harm MW, Seidenberg MS. 2004. Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. Psychol Rev 111:662-720.

Helenius P, Uutela K, Hari R. 1999. Auditory stream segregation in dyslexic adults. Brain 122 (Pt 5):907-913.

Henderson VW. 1986. Anatomy of posterior pathways in reading: A reassessment. Brain Lang 29:199-233.

Hinke RM, Hu X, Stillman AE, Kim SG, Merkle H, Salmi R, Ugurbil K. 1993. Functional magnetic resonance imaging of Broca's area during internal speech. Neuroreport 4:675-678 675-678.

Huettel SA, Song AW, McCarthy G. 2005. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. J Neurosci 25:3304-3311.

Huttenlocher P, Dabholkar A. 1997. Regional differences in synaptogenesis in human cerebral cortex. The Journal of Comparative Neurology.

Jared D. 1997. Evidence that strategy effects in word naming reflect changes in output timing rather than changes in processing route. J Exp Psychol Learn Mem Cogn 23:1424-1438.

Jobard G, Crivello F, Tzourio-Mazoyer N. 2003. Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. Neuroimage 20:693-712.

Joubert S, Beauregard M, Walter N, Bourgouin P, Beaudoin G, Leroux JM, Karama S, Lecours AR. 2004. Neural correlates of lexical and sublexical processes in reading. Brain Lang 89:9-20.

Kail R. 1991. Development of processing speed in childhood and adolescence. Adv Child Dev Behav 23:151-185.

Kail R. 1991. Developmental change in speed of processing during childhood and adolescence. Psychol Bull 109:490-501.

Kail RV, Miller CA. 2006. Developmental change in processing speed: Domain specificity and stability during childhood and adolescence. Journal of Cognition and Development 7:119-137.

Kello CT, Plaut DC. 2003. Strategic control over rate of processing in word reading: A computational investigation. Journal of Memory & Language 48:207-232.

Kherif F, Josse G, Price CJ. 2011. Automatic top-down processing explains common left occipito-temporal responses to visual words and objects. Cereb Cortex 21:103-114.

Kinoshita S, Lupker SJ. 2003. Priming and attentional control of lexical and sublexical pathways in naming: A reevaluation. J Exp Psychol Learn Mem Cogn 29:405-415.

Kopp CB. 1982. Antecedents of self-regulation: A developmental perspective. Dev Psychol 18:199-214.

Koyama MS, Di Martino A, Kelly C, Jutagir DR, Sunshine J, Schwartz SJ, Castellanos FX, Milham MP. 2013. Cortical signatures of dyslexia and remediation: an intrinsic functional connectivity approach. PLoS ONE 8:e55454.

Kutas M, Hillyard SA. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207:203-205.

Laasonen M, Tomma-Halme J, Lahti-Nuuttila P, Service E, Virsu V. 2000. Rate of information segregation in developmentally dyslexic children. Brain Lang 75:66-81.

LaBerge D, Samuels J. 1974. Toward a theory of automatic information processing in reading. Cognit Psychol 6:293-323.

Lerner JW. 1989. Educational interventions in learning disabilities. J Am Acad Child Adolesc Psychiatry 28:326-331.

Liberman IY, Shankweiler D, Liberman AM. 1989. Phonology and reading disability: Solving the reading puzzle. In: Shankweiler D, Liberman IY, editors. International academy for research in learning disabilities monograph series Ann Arbor, MI: University of Michigan Press p 1-33.

Liu J, Li J, Zhang H, Rieth CA, Huber DE, Li W, Lee K, Tian J. 2010. Neural correlates of top-down letter processing. Neuropsychologia 48:636-641.

Logan GD. 1997. Automaticity and reading: Perspectives from the instance theory of automatization. Reading and Writing Quarterly: Overcoming Learning Difficulties 13:123-146.

Luck SJ, Vogel EK, Shapiro KL. 1996. Word meanings can be accessed but not reported during the attentional blink. Nature 383:616-618.

Luna B, Garver KE, Urban TA, Lazar NA, Sweeney JA. 2004. Maturation of cognitive processes from late childhood to adulthood. Child Dev 75:1357-1372.

Luna B, Thulborn K, Munoz D, Merriam E. 2001. Maturation of Widely Distributed Brain Function Subserves Cognitive Development. Neuroimage.

Lupker SJ, Brown P, Colombo L. 1997. Strategic control in a naming task: Changing routes or changing deadlines? J Exp Psychol Learn Mem Cogn 23:570-590.

MacDonald AW, Cohen JD, Stenger VA, Carter CS. 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288:1835-1838.

MacLeod CM. 1991. Half a century of research on the Stroop effect: an integrative review. Psychol Bull 109:163-203.

McAvoy MP, Ollinger JM, Buckner RL. 2001. Cluster size thresholds for assessment of significant activation in fMRI. Neuroimage 13:S198.

McCandliss BD, Noble KG. 2003. The development of reading impairment: a cognitive neuroscience model. Ment Retard Dev Disabil Res Rev 9:196-204.

McCann RS, Besner D. 1987. Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. Journal of Experimental Psychology: Human Perception 13:14-24.

McCann RS, Remington RW, Van Selst M. 2000. A dual-task investigation of automaticity in visual word processing. J Exp Psychol Hum Percept Perform 26:1352-1370.

McCarthy G, Blamire AM, Rothman DL, Gruetter R, Shulman RG. 1993. Echo-planar magnetic resonance imaging studies of frontal cortex activation during word generation in humans. Proceedings of the National Academy of Science 90:4952-4956.

McDermott KB, Petersen SE, Watson JM, Ojemann JG. 2003. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. Neuropsychologia 41:293-303.

McGrath LM, Pennington BF, Shanahan MA, Santerre-Lemmon LE, Barnard HD, Willcutt EG, Defries JC, Olson RK. 2011. A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: searching for shared cognitive deficits. J Child Psychol Psychiatry 52:547-557.

Mechelli A, Gorno-Tempini ML, Price CJ. 2003. Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. J Cogn Neurosci 15:260-271.

Mechelli A, Josephs O, Lambon Ralph MA, McClelland JL, Price CJ. 2007. Dissociating stimulus-driven semantic and phonological effect during reading and naming. Hum Brain Mapp 28:205-217.

Meiran N. 1996. Reconfiguration of processing mode prior to task performance. Journal of Experimental Psychology: Human Learning and Memory 22:1423-1442.

Monsell S, Doyle MC, Haggard PN. 1989. Effects of frequency on visual word recognition tasks: Where are they? Journal of Experimental Psychology: General 118:43-71.

Monsell S, Patterson K, Graham A, Hughes CH, Milroy R. 1992. Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. Journal of Experimental Psychology: Learning, Memory and Cognition 18:452-467.

Morton J. 1969. Interaction of information in word recognition. Psychol Rev 76:165-178.

Nagy WE, Anderson RC. 1995. Metalinguistic awareness and literacy acquisition in different languages. In. Technical Report No. 618. University of Illinois at Urbana-Champaign: Center for the Study of Reading.

Nakamura K, Dehaene S, Jobert A, Le Bihan D, Kouider S. 2007. Task-specific change of unconscious neural priming in the cerebral language network. Proc Natl Acad Sci U S A 104:19643-19648.

Neely JH. 1977. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. J Experimental Psychology: General 106:226-254.

Nelles JL, Lugar HM, Coalson RS, Miezin FM, Petersen SE, Schlaggar BL. 2003. An automated method for extracting response latencies of subject vocalizations in event-related fMRI experiments. Neuroimage 20:1865-1871.

Nelson SM, Dosenbach NU, Cohen AL, Wheeler ME, Schlaggar BL, Petersen SE. 2010. Role of the anterior insula in task-level control and focal attention. Brain Struct Funct 214:669-680.

Neta M, Schlaggar BL, Petersen SE. Separable responses to error, ambiguity, and reaction time in cingulo-opercular task control regions. Under revision.

Nickels L, Biedermann B, Coltheart M, Saunders S, Tree JJ. 2008. Computational modelling of phonological dyslexia: how does the DRC model fare? Cogn Neuropsychol 25:165-193.

Nomura EM, Gratton C, Visser RM, Kayser A, Perez F, D'Esposito M. 2010. Double dissociation of two cognitive control networks in patients with focal brain lesions. Proc Natl Acad Sci U S A 107:12017-12022.

Norman DA, Shallice T. 1986. Attention to action: Willed and automatic control of behavior. In. Consciousness and Self-regulation Plenum Press p 1-18.

Norris D. 1994. A quantitative multiple-levels model of reading aloud. Special Section: Modeling visual word recognition. Journal of Experimental Psychology: Human Perception and Performance 20:1212-1232.

Norris D, Kinoshita S. 2012. Reading through a noisy channel: why there's nothing special about the perception of orthography. Psychol Rev 119:517-545.

Ollinger JM, Shulman GL, Corbetta M. 2001. Separating processes within a trial in event-related functional MRI I. The method. Neuroimage 13:210-217.

Paap KR, Noel RW. 1991. Dual route models of print to sound: Still a good horse race. Psychol Res 53:13-24.

Pachella RG. 1974. The interpretation of reaction time in information-processing research. In: Kantowitz BH, editor. Human information processing: Tutorials in performance and cognition. Potomac, MD: Erlbaum p 41-82.

Paloyelis Y, Rijsdijk F, Wood AC, Asherson P, Kuntsi J. 2010. The genetic association between ADHD symptoms and reading difficulties: the role of inattentiveness and IQ. J Abnorm Child Psychol 38:1083-1095.

Papanicolaou AC, Simos PG, Basile LFA. 1998. Applications of magnetoencephalography in linguistic research. In: Stemmer B, Whitaker HA, editors. The handbook of neurolinguistics New York: Academic Press p 143-158.

Pardo JV, Pardo PJ, Janer KW, Raichle ME. 1990. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. Proc Natl Acad Sci U S A 87:256-259.

Perfetti CA. 1992. The representation problem in reading acquisition. In: Gough PB, Ehri LC, Treiman R, editors. Reading Acquisition Hillsdale, NJ: Erlbaum p 145-174.

Perrone-Bertolotti M, Kujala J, Vidal JR, Hamame CM, Ossandon T, Bertrand O, Minotti L, Kahane P, Jerbi K, Lachaux JP. 2012. How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. J Neurosci 32:17554-17562.

Perry C, Ziegler JC, Zorzi M. 2007. Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. Psychol Rev 114:273-315.

Perry C, Ziegler JC, Zorzi M. 2013. A Computational and Empirical Investigation of Graphemes in Reading. Cogn Sci.

Petersen SE, Dubis JW. 2012. The mixed block/event-related design. Neuroimage 62:1177-1184.

Petersen SE, Fiez JA. 1993. The processing of single words studied with positron emission tomography. Annu Rev Neurosci 16:509-530.

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331:585-589.

Petersen SE, Fox PT, Snyder AZ, Raichle ME. 1990. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. Science 249:1041-1044.

Petersen SE, Posner MI. 2012. The attention system of the human brain: 20 years after. Annu Rev Neurosci 35:73-89.

Peterson BS, Skudlarski P, Gatenby JC, Zhang H, Anderson AW, Gore JC. 1999. An fMRI study of Stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. Biol Psychiatry 45:1237-1258.

Peterson RL, Pennington BF. 2012. Developmental dyslexia. Lancet 379:1997-2007.

Petrides M, Alivisatos B, Evans AC. 1995. Functional activation of the human ventrolateral frontal cortex during mnemonic retrieval of verbal information. Proc Natl Acad Sci U S A 92:5803-5807.

Peyrin C, Lallier M, Demonet JF, Pernet C, Baciu M, Le Bas JF, Valdois S. 2012. Neural dissociation of phonological and visual attention span disorders in developmental dyslexia: FMRI evidence from two case reports. Brain Lang 120:381-394.

Platt ML, Glimcher PW. 1999. Neural correlates of decision variables in parietal cortex. Nature 400:233-238.

Plaut DC, McClelland JL, Seidenberg MS, Patterson K. 1996. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychol Rev 103:56-115.

Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME. 2007. Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. J Neurosci 27:11912-11924.

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 10:15-35.

Posner MI, Petersen SE. 1990. The attention system of the human brain. Annu Rev Neurosci 13:25-42.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142-2154.

Power JD, Cohen AL, Nelson SM, Vogel AC, Church JA, Barnes KA, Wig GS, Laumann TO, Meizin FM, Schlaggar BL, Petersen SE. 2011. Functional network organization in the human brain. Neuron 72:665-678.

Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. In preparation.

Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage 62:816-847.

Price CJ, Devlin JT. 2003. The myth of the visual word form area. Neuroimage 19:473-481.

Price CJ, Devlin JT. 2011. The interactive account of ventral occipitotemporal contributions to reading. Trends Cogn Sci 15:246-253.

Pritchard SC, Coltheart M, Palethorpe S, Castles A. 2012. Nonword reading: comparing dual-route cascaded and connectionist dual-process models with human data. J Exp Psychol Hum Percept Perform 38:1268-1288.

Protopapas A, Archonti A, Skaloumbakas C. 2007. Reading ability is negatively related to Stroop interference. Cogn Psychol 54:251-282.

Pugh KR, Mencl WE, Jenner AR, Katz L, Frost SJ, Lee JR, Shaywitz SE, Shaywitz BA. 2001. Neurobiological studies of reading and reading disability. J Commun Disord 34:479-492.

Pugh KR, Rexer K. 1994. Evidence of flexible coding in visual word recognition. Journal of Experimental Psychology: Human Perception and Performance 20:807-825.

Pugh KR, Shaywitz BA, Shaywitz SE, Constable RT, Skudlarski P, Fulbright RK, Bronen RA, Shankweiler DP, Katz L, Fletcher JM, Gore JC. 1996. Cerebral organization of component processes in reading. Brain 119 (Pt 4):1221-1238.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. Proc Natl Acad Sci U S A 98:676-682.

Ramus F, Ahissar M. 2012. Developmental dyslexia: the difficulties of interpreting poor performance, and the importance of normal performance. Cogn Neuropsychol 29:104-122.

Ramus F, Rosen S, Dakin SC, Day BL, Castellote JM, White S, Frith U. 2003. Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. Brain 126:841-865.

Rastle K. 2012. Rethinking phonological theories of reading. Behav Brain Sci 35:303-304.

Rastle K, Coltheart M. 1999. Serial and strategic effects in reading aloud. J Exp Psychol Hum Percept Perform 25:482-503.

Ratcliff R. 1979. Group reaction time distributions and an analysis of distribution statistics. Psychol Bull 86:446-461.

Ratcliff R, Murdock BB, Jr. 1976. Retrieval processes in recognition memory. Psychol Bull 86:446-461.

Reicher GM. 1969. Perceptual recognition as a function of meaningfulness of stimulus material. J Exp Psychol 81:275-280.

Reynolds M, Besner D. 2005. Contextual control over lexical and sublexical routines when reading english aloud. Psychon Bull Rev 12:113-118.

Reynolds M, Besner D. 2006. Reading aloud is not automatic: processing capacity is required to generate a phonological code from print. J Exp Psychol Hum Percept Perform 32:1303-1323.

Reynolds M, Mulatti C, Besner D. 2012. Reading nonwords aloud: evidence for dynamic control in skilled readers. Psychon Bull Rev 19:1135-1141.

Roskies AL, Fiez JA, Balota DA, Raichle ME, Petersen SE. 2001. Task-dependent modulation of regions in left frontal cortex during semantic processing. J Cogn Neurosci 13:829-843.

Rossmeissl PG, Theios J. 1982. Identification and pronunciation effects in a verbal reaction time task for words, pseudowords, and letters. Mem Cognit 10:443-450.

Rueda MR, Fan J, McCandliss BD, Halparin JD, Gruber DB, Lercari LP, Posner MI. 2004. Development of attentional networks in childhood. Neuropsychologia 42:1029-1040.

Rumsey JM, Horwitz B, Donohue BC, Nace K, Maisog JM, Andreason P. 1997. Phonological and orthographic components of word recognition: A PET-rCBF study. Brain 120:739-759.

Rushworth MFS, Paus T, Sipila PK. 2001. Attention systems and the organization of the human parietal cortex. J Neurosci 21:5262-5271.

Ruz M, Nobre AC. 2008. Attention modulates initial stages of visual word processing. J Cogn Neurosci 20:1727-1736.

Ruz M, Nobre AC. 2008. Dissociable top-down anticipatory neural states for different linguistic dimensions. Neuropsychologia 46:1151-1160.

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH. 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64:240-256.

Schilling HH, Rayner K, Chumbley JI. 1998. Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. Mem Cognit 26:1270-1281.

Schlaggar BL, McCandliss BD. 2007. Development of neural systems for reading. Annu Rev Neurosci 30:475-503.

Schneider W, Shiffrin RM. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. Psychol Rev 84:1-53.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349-2356.

Seghier ML. 2013. The angular gyrus: multiple functions and multiple subdivisions. Neuroscientist 19:43-61.

Seghier ML, Price CJ. 2012. Functional heterogeneity within the default network during semantic processing and speech production. Front Psychol 3:281.

Seidenberg MS. 1985. The time course of information activation and utilization in visual word recognition. In: Mackinnon GE, Waller TG, editors. Reading Research: Advances in Theory and Practice New York: Academic Press, Inc p 199-251.

Seidenberg MS, McClelland JL. 1989. A distributed, developmental model of word recognition and naming. Psychol Rev 96:523-568.

Sestieri C, Shulman GL, Corbetta M. 2010. Attention to memory and the environment: functional specialization and dynamic competition in human posterior parietal cortex. J Neurosci 30:8445-8456.

Share DL, Stanovich KE. 1995. Cognitive processes in early reading development: Accomodating individual differences into a model of acquisition. Issues in Education 1:1-57.

Shaywitz SE, Shaywitz BA. 2005. Dyslexia (specific reading disability). Biol Psychiatry 57:1301-1309.

Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE. 2013. Statistical improvements in fMRI analyses produced by censoring high motion datapoints. Hum Brain Mapp. *In press.*

Siegel LS. 2006. Perspectives on dyslexia. Paediatr Child Health 11:581-587.

Smith EE. 1968. Choice reaction time: an analysis of the major theoretical positions. Psychol Bull 69:77-110.

Spieler DH, Balota DA. 1997. Bringing computational models of word naming down to the item level. Psychol Sci 8:411-416.

Spieler DH, Balota DA, Faust ME. 1996. Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. J Exp Psychol Hum Percept Perform 22:461-479.

Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL. 2010. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. Neuroimage 53:303-317.

Sprenger-Charolles L, Siegel LS, Bonnet P. 1998. Reading and spelling acquisition in French: the role of phonological mediation and orthographic factors. J Exp Child Psychol 68:134-165.

Stanovich KE. 1990. Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. Dev Rev 10:72-100.

Stanovich KE. 1993. The language code: Issues in word recognition. In: Yussen SR, Smith MC, editors. Reading across the life span New York: Springer-Verlag p 111-135.

Stein J, Walsh V. 1997. To see but not to read; the magnocellular theory of dyslexia. Trends Neurosci 20:147-152.

Sternberg S. 1969. The discovery of processing stages: extensions of Donders' method. In. Acta Psychologica   North Holland Publishing p 276-315.

Stroop JR. 1935. Studies of interference in serial verbal reactions. J Exp Psychol 18:643-662.

Talairach J, Tournoux P. 1988. Co-Planar Stereotaxic Atlas of the Human Brain. New York: Thieme Medical Publishers, Inc.

Tan LH, Spinks JA, Eden GF, Perfetti CA, Siok WT. 2005. Reading depends on writing, in Chinese. Proc Natl Acad Sci U S A 102:8781-8785.

Tanenhaus MK, Flanigan HP, Seidenberg MS. 1980. Orthographic and phonological activation in auditory and visual word recognition. Mem Cognit 8:513-520.

Taylor JS, Rastle K, Davis MH. 2012. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. Psychol Bull.

Theios J, Muise JG. 1977. The word identification process in reading. In: Castellan NJ, Jr., Pisoni DB, Potts GR, editors. Cognitive Theory Hillsdale, N.J.: Erlbaum.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. 1997. Role of the left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proceedings of the National Academy of Sciences, USA 94:14792-14797.

Treiman R, Kessler B. 1995. In defense of an onset-rime syllable structure for English. Lang Speech 38 (Pt 2):127-142.

Turkeltaub PE, Gareau L, Flowers DL, Zeffiro TA, Eden GF. 2003. Development of neural mechanisms for reading. Nat Neurosci 6:767-773.

Turkeltaub PE, Goldberg EM, Postman-Caucheteux WA, Palovcak M, Quinn C, Cantor C, Coslett HB. 2013. Alexia due to ischemic stroke of the visual word form area. Neurocase.

Twomey T, Kawabata Duncan KJ, Price CJ, Devlin JT. 2011. Top-down modulation of ventral occipito-temporal responses during visual word recognition. Neuroimage 55:1242-1251.

Valdois S, Bosse ML, Tainturier MJ. 2004. The cognitive deficits responsible for developmental dyslexia: review of evidence for a selective visual attentional disorder. Dyslexia 10:339-363.

Van Essen DC, Dickson J, Harwell J, Hanlon D, Anderson CH, Drury HA. 2001. An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assoc 41:1359-1378. See also http://brainmap.wustl.edu/caret.

Vidyasagar TR, Pammer K. 2010. Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. Trends Cogn Sci 14:57-63.

Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL. 2008. Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. J Neurophysiol 100:3328-3342.

Visser TA, Besner D. 2001. On the dominance of whole-word knowledge in reading aloud. Psychon Bull Rev 8:560-567.

Vogel AC, Miezin FM, Petersen SE, Schlaggar BL. 2012. The putative visual word form area is functionally connected to the dorsal attention network. Cereb Cortex 22:537-549.

Vogel AC, Petersen SE, Schlaggar BL. 2012. The left occipitotemporal cortex does not show preferential activity for words. Cereb Cortex 22:2715-2732.

Vogel AC, Petersen SE, Schlaggar BL. 2012. Matching is not naming: A direct comparison of lexical manipulations in explicit and implicit reading tasks. Hum Brain Mapp.

Vogel EK, Woodman GF, Luck SJ. 2005. Pushing around the locus of selection: evidence for the flexible-selection hypothesis. J Cogn Neurosci 17:1907-1922.

Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA. 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31:329-338.

Wagner RK, Torgesen JK, Rashotte CA. 1999. Comprehensive test of phonological processing (CTOPP). Austin, TX: Pro-Ed.

Wechsler D. 1999. Wechsler abbreviated scale of intelligence. San Antonio: The Psychological Corporation.

Weissman DH, Roberts KC, Visscher KM, Woldorff MG. 2006. The neural bases of momentary lapses in attention. Nat Neurosci 9:971-978.

Wickens CD. 1974. Temporal limits of human information processing: A developmental study. Psychol Bull 81:739-755.

Wilkinson D, Halligan P. 2004. Opinion: The relevance of behavioural measures for functional-imaging studies of cognition. Nat Rev Neurosci 5:67-73.

Willcutt EG, Betjemann RS, McGrath LM, Chhabildas NA, Olson RK, DeFries JC, Pennington BF. 2010. Etiology and neuropsychology of comorbidity between RD and ADHD: the case for multiple-deficit models. Cortex 46:1345-1361.

Willcutt EG, Pennington BF. 2000. Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. J Learn Disabil 33:179-191.

Woodcock RW, Johnson MB. 2002. Woodcock-Johnson-Revised Tests of Achievement. In. Itasca, IL: Riverside Publishing.

Yap MJ, Balota DA, Sibley DE, Ratcliff R. 2012. Individual differences in visual word recognition: insights from the English Lexicon Project. J Exp Psychol Hum Percept Perform 38:53-79.

Yap MJ, Tse CS, Balota DA. 2009. Visual word recognition of multisyllabic words. J Mem and Lang 61:303.

Yarkoni T, Barch DM, Gray JR, Conturo TE, Braver TS. 2009. BOLD correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fMRI analysis. PLoS ONE 4:e4257.

Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125-1165.

Yoncheva YN, Zevin JD, Maurer U, McCandliss BD. 2010. Auditory selective attention to speech modulates activity in the visual word form area. Cereb Cortex 20:622-632.

Zelazo PD, Craik FI, Booth L. 2004. Executive function across the life span. Acta Psychol (Amst) 115:167-183.

Zeno SM, Ivens SH, Millard RT, Duvvuri R. 1995. The educator's word frequency guide. Brewster, NY: Touchstone Applied Science Associates.

Zevin JD, Balota DA. 2000. Priming and attentional control of lexical and sublexical pathways during naming. J Exp Psychol Learn Mem Cogn 26:121-135.

Zhao J, Liu J, Li J, Liang J, Feng L, Ai L, Lee K, Tian J. 2011. Intrinsically organized network for word processing during the resting state. Neurosci Lett 487:27-31.

Zoccolotti P, De Luca M, Di Pace E, Gasperini F, Judica A, Spinelli D. 2005. Word length effect in early reading and in developmental dyslexia. Brain Lang 93:369-373.

Zorzi M, Houghton G, Butterworth B. 1998. Two routes or one in reading aloud? A connectionist dual-process model. J Exp Psychol Hum Percept Perform 24:1131-1161.