IDIOLECT

INTRODUCTION

"Idiolect" refers to an individual's unique variety and/or use of language, from the level of the phoneme to the level of discourse. This meaning is reflected in the etymology of the word: the two morphemes *idio-* and *-lect*. *Idio-* is of Greek origin, and means "own, personal, private, peculiar, separate and distinct," while *-lect* refers to a "social variety of a language." The theory holds, therefore, that no two people who share a common language have exactly the same linguistic repertoire. In the same way that the variation exhibited in a person's language production is influenced by their dialect(s), sociolect(s) and by register, so too is it influenced by their personal, idiosyncratic, often habitual linguistic preferences—their idiolect. A person's idiolect is all encompassing in that it includes linguistic features related to dialect and sociolect, for example, while also being influenced by a wide range of other sources of variation, such as their life experiences; language encounters; what they have read and listened to; where they have been schooled; jobs they have had; their favorite hobbies and pastimes; and their parents, friends, and teachers. An idiolect, therefore, is not stable in its entirety. While some elements may persist throughout a person's life, others may drop out of favor, while new patterns, preferences, and features may be acquired over time. Despite being generally accepted in linguistics, the concept of idiolect has received relatively little rigorous or systematic research attention. However, since its introduction in linguistics in the late nineteenth century, the notion of linguistic individuality or the role of the individual in language, if not explicitly the term "idiolect," has sporadically been the focus of discussion across a range of disciplines in linguistics. This includes debates over whether idiolect resides in the overall linguistic system of an individual or in their patterns of usage, or whether the individual plays a role in language change. In some fields, the advent of large collections of texts has facilitated a testing of the theory of idiolect. Nevertheless, the concept of idiolect is familiar and mysterious in equal measure in the linguistics literature. It is a term that has a passing mention or glossary entry in most introductory textbooks in linguistics,

yet it is a theory that is not easily observable or measurable, and for which there is little agreement and even less empirical evidence.

ORIGINS OF THE TERM "IDIOLECT"

Notions of linguistic individuality can be traced back to the earliest formalized discussions of language and linguistics. The most notable of such early studies is Paul 1888, in which all language originates in an individual mind, and therefore the individual should be the focus of scrutiny. The role of the individual is also highlighted in Saussure 1983 in the distinction between *langue* and *parole*. Saussure argues that the execution of language is always individual, and refers to this as *speech (parole)*, therefore distinguishing the language system itself (*langue*) from speech actually produced by individuals. In doing so, the work highlights that we can distinguish what is general and social from what is individual. Sapir 1927 presents individual linguistic variation as a window through which to analyze personality traits of speakers. This work seeks to disambiguate elements of speech that are social norms and those that are reflective of individual expression and personality. Contemporary to that work, Bloomfield 1933, also draws a distinction between the speech community as a collective and the individuals of which it is comprised. Bloomfield notes that one of the difficulties in determining which people belong to the same speech community is that no two people speak exactly alike. The term "idiolect" is widely accepted to have been first used in Bloch 1948, to refer to "the totality of the possible utterances of one speaker at one time in using language to interact with one other speaker." The article specifies that an idiolect is not merely what a speaker says at one time: it is everything that he *could* say in a given language at that specific time. Since Bloch 1948, definitions of idiolect have varied, as some refer to the language system available to the individual and others to the samples of language actually produced by the individual. Hockett 1958, for example, is similar to Bloch in defining idiolect as "the totality of speech habits of a single person at a given time," distinguishing habits from observable behavior. On the other hand, Martinet 1961 defines idiolect as "the language *as spoken* by a single individual" (emphasis not in the original).

Bloch, Bernard. 1948. A set of postulates for phonemic analysis. *Language* 24.1: 3–46.

 The work in which the term "idiolect" was coined. In this article, which is primarily focused on characteristics of spoken language, Bloch lays the theoretical foundations for the concept of

idiolect that have been the basis for discussions in the decades since. It is widely cited, given its debuting of the term.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart, and Winston.
One of the most important works in linguistics published in the twentieth century. Linguistic individuality receives a passing mention as a problem for defining and identifying homogenous speech communities. However, this mention demonstrates how "the individual" came to be discussed in contrast with "the group."

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.
An important work in linguistics that aims to present "the generally accepted facts and principles" of the field of linguistics. Idiolect is most prominent in chapter 38, in which it is discussed alongside concepts of "dialect" and "language."

Martinet, A. 1961. *A functional view of language*. Oxford: Oxford Univ. Press.
Although idiolect only receives brief attention in this book between pages 104 and 108, it is worth noting as Martinet discusses idiolect from a functional perspective of language, and therefore offers an interpretation of idiolect as relating to language *production* rather than language *systems*, which differs from others writing at the time.

Paul, Hermann. 1888. *Principles of the history of language*. Translated by H. A. Strong. London: Swan, Sonnenschien, Lowrey.
One of the earliest and most important works in linguistics, which puts the individual and the individual mind at the center of language study. Although the ideas in this book foreshadow many of the later discussions of the individual in linguistics, following the work of Saussure and subsequent directions in grammatical theory and variationist sociolinguistics, Paul's work and its focus on the individual has been, to some extent, buried.

Sapir, Edward. 1927. Speech as a personality trait. *American Journal of Sociology* 32.6: 892–905.

  A seminal work in which Sapir separates out notions of group norms and linguistic individuality. In the same way as Paul 1888, Sapir traverses the intersections between psychology and language. The individual features throughout this work.

Saussure, Ferdinand de. 1983. *Course in general linguistics*. Edited by Charles Bally and Albert Sechehaye. Translated and annotated by Roy Harris. London: Duckworth. [ISBN: 9780812697063]

  The original version of this work, written in French and published in 1916, is widely regarded as the foundational cornerstone of modern-day formal linguistics. The distinction drawn between *language* and *speech* has defined the directions and trajectories of different fields of linguistics for over a century. This work is important for idiolect as it argues that it is only through individuals that language systems are manifest and observable.

IDIOLECT IN SOCIOLINGUISTICS

Most introductory textbooks in sociolinguistics or language variation and change will include some mention or glossary entry for the term "idiolect." However, the concept has not always had such a recognized status in variationist studies. The role of the individual in language variation and change was dismissed in Weinreich 1954, and then more famously in Weinreich, et al. 1968, which argues that language change occurs at the level of the speech community, rather than single individuals. Labov 1972 famously reiterates this argument, emphasizing that language is an instrument used by members of a community to communicate with each other, and therefore "idiosyncratic habits" are not part of language so conceived. These are important works in the history of idiolect because they served to background the concept in sociolinguistics for decades. There are some exceptions to the dominant paradigm in sociolinguistics which followed, however. For example, unlike Labov, Lieb 1993 considers idiolects as central to the composition of language, and defines an idiolect as a "means of communication for some person during a certain time interval." In turn, Lieb 1993 argues that language is construed as a set of such means of communications. Milroy 1997 argues that linguistic change is initiated by individual speakers,

and such changes are a product of speaker innovation and the passing on from one speaker to others through interaction. This argument is central to other studies, such as Auer and Hinskens 2005, which focuses on interpersonal linguistic accommodation between individuals and the part it plays in long term structural change. The individual-in-interaction is also the foundation for Kuhl 2003, which argues that language change is a product of idiolectal contact between speakers. Taking a similar stance, the individual features prominently in Mufwene 2001 on language evolution. The most comprehensive discussion of the so-called linguistic individual is Johnstone 1996, which branches across discourse analysis, pragmatics, and sociolinguistics. Although the word "idiolect" appears only twice in the book, speaker idiosyncrasies and the construction of the individual voice is examined across various parameters of language use and discourse types. Here, Johnstone 1996 emphasizes the role of the individual in the performance and construction of identity and personae, a concept that subsequently characterized Third Wave sociolinguistics. More recently, Johnstone 2009 revisits a case study introduced in the 1996 book—that of the individual speech style of U.S. political figure Barbara Jordan.

Auer, Peter, and Frans Hinskens. 2005. The role of interpersonal accommodation in a theory of language change. In *Dialect change. The convergence and divergence of dialects in contemporary societies.* Edited by Peter Auer, Frans Hinskens, and Paul Kerswill, 335–357. Cambridge, UK: Cambridge Univ. Press.
This chapter sets out to empirically test the notion that structural language change is intrinsically linked to linguistic accommodation in face-to-face verbal communication. Such change-by-accommodation elevates the role of the individual in language change.

Johnstone, Barbara. 1996. *The linguistic individual: Self-expression in language and linguistics*. Oxford: Oxford Univ. Press. [ISBN: 9780195101843]
One of only very few book-length treatments of the linguistics of the individual speaker and therefore a very important contribution to the field. The book describes the status of the individual in linguistics, before examining individual linguistic variation and expression across various contexts. The book ends with a recommendation for linguists to "bring individuals back into the picture" of language study.

Johnstone, Barbara. 2009. Stance, style, and the linguistic individual. In *Stance: Sociolinguistic perspectives*. Edited by Alexander Jaffe, 29–52. Oxford: Oxford Univ. Press. [ISBN: 9780195331646]

Taking a discourse analysis approach, Johnstone revisits a case study of Barbara Jordan's idiolect, presenting an analysis that demonstrates how patterns of stance-taking constitute a personal linguistic style, which in turn is influenced by a speaker's life history and sociolinguistic environment.

Kuhl, Joseph W. 2003. *The idiolect, chaos, and language custom far from equilibrium: Conversations in Morocco[https://getd.libs.uga.edu/pdfs/kuhl_joe_w_200308_phd.pdf]*. PhD thesis, Univ. of Georgia, Athens.

An important but generally under-cited contribution to the idiolect debate in sociolinguistics. This doctoral dissertation lays out the theoretical foundations of idiolect first proposed in Paul 1888 (cited in *Origins of the Term "Idiolect"*) and offers a history of the concept in linguistics and linguistic theory, before analyzing conversations in which interlocutors' idiolects are very different, or "far-from-equilibrium".

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: Univ. of Pennsylvania. [ISBN: 9780812210521]

This book is a seminal work in the field of variationist sociolinguistics. The book reports on the social stratification of linguistic variables in Martha's Vineyard and in New York, and addresses the theoretical underpinnings of language change. This book is important for idiolect as it eschewed idiosyncratic variation as being of little relevance to language change. With Labov's and the book's popularity came the backgrounding of idiolect in sociolinguistic study.

Lieb, Hans-Heinrich. 1993. *Linguistic variables: Towards a unified theory of linguistic variation*. Amsterdam: John Benjamins. [ISBN: 9789027236111]

"Unified" in the title of this book refers to Lieb's efforts in developing a theoretical framework of language variation comprising approaches not only from sociolinguistics and dialectology, but more broadly across areas such as generative grammar, psycholinguistics, and stylistics.

Lieb's approach hinges on the concept of the linguistic variable, and the notion of idiolect features heavily throughout.

Milroy, James. 1997. Internal vs. external motivations for linguistic change. *Multilingua* 16.4: 311–323.

A relatively short article that makes no direct mention of the term "idiolect." However, it is notable in sociolinguistics as it emphasizes the role of the individual and interactions between individuals in processes of language change and language contact, contrary to the ruling paradigm in the field after Weinreich, et al. 1968 and Labov 1972.

Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge, UK: Cambridge Univ. Press. [ISBN: 9780521791380]

This book unambiguously attributes long-term language change and evolution to the contact between individual idiolects. As the title suggests, metaphors from ecology are used throughout to explain language, dialect, and idiolect contact.

Weinreich, Uriel. 1954. Is structural dialectology possible? *Word* 10.2–3: 388–400.

Aims to combine the fields of structural and dialectological studies in linguistics. Idiolect is described as a reduction of language to its extreme and to "absurdity," and the study of individual idiolects as is labeled inexhaustible and hardly worth the effort. This is an important argument that is later made in more detail by Weinreich, et al. 1968, and which shaped sociolinguistics for most of the twentieth century.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for historical linguistics: A symposium*. Edited by Winfred Philipp Lehmann and Yakov Malkiel, 98–195. Austin: Univ. of Texas Press.

A seminal paper in sociolinguistics that introduces and discusses major concepts such as orderly heterogeneity and the actuation problem, and served as the foundation of language variation and change studies in the decades since. This paper is important in the context of

idiolect in that it did more than most to background the study of the individual in favor of focusing on the speech community.

IDIOLECT AND GRAMMAR

Although seldom referred to explicitly as "idiolect," the language of the individual features across a range of important works in grammatical and lexical theory from a breadth of different perspectives. From the generative viewpoint, "I-Language" was first introduced in Chomsky 1986 to refer to the internalized language faculty in the mind of the individual speaker, as opposed to "E-Language" or externalized language, which exists outside of the mind and is shared socially. For Chomsky, I-Language was the object of linguistic study. The concepts of I-Language and E-Language are further discussed in relation to individuation and idiolect in Ludlow 2011, in which the parametric states of the former are said to vary from individual to individual. In work emerging from the Systemic Functional Linguistics (SFL) tradition, the individual comes to the fore in various ways. In a language teaching context, Halliday, et al. 1964 discusses idiolect in relation to dialect and register, and conceptualizes a person's idiolect as heterogeneous, changeable, unique, and recognizable by its grammatical and lexical characteristics. Elsewhere within SFL, though primarily in relation to semantics rather than grammar, the individual features in the concept of "code" in Hasan 2002, the notion of "personalized meaning potential" discussed at length in Matthiessen 2009, and the process of "individuation" as described in Martin 2009. Somewhat inevitably, the individual plays a central role in usage-based theories of grammar, and Schmid and Mantlik 2015 investigates the extent to which a particular grammatical structure is "entrenched" in the minds of 83 individual historical authors. Similarly, in the theory of lexical priming, Hoey 2005 emphasizes the personal and idiolectal nature of collocations. In contrast, Harris 1998 criticizes and rejects many of the tenets underlying the "idiolect myth."

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. Convergence. New York: Praeger. [ISBN: 9780275917616]

The famous, well-cited book on the philosophy of language and philosophy of mind in which the distinction between I-Language and E-Language was first drawn.

Halliday M. A. K., Angus McIntosh, and Peter Strevens. 1964. *The linguistic sciences and language teaching*. London: Longman.

  The term "idiolect" does not feature much in Halliday's work. This book is split into two parts, the first called "The Linguistic Sciences" and the second called "The Linguistic Sciences in Relation to Language Teaching and Language Learning." Chapter 5 of part 1 focuses on the users and uses of language, and it is here where the concept of idiolect gets a few pages of attention.

Harris, Roy. 1998. *Introduction to integrational linguistics*. Oxford: Pergamon. [ISBN: 9780080433646]

  An introduction to integrationism as a new approach to general linguistics departing from structuralism and generativism. Harris challenges the concept of idiolect in terms of the assumption that individuals constantly speak in a characteristic way, and its implications for communication between two people who are considered to have uniquely different "codes," and the passing on of languages through generations.

Hasan, Ruqaiya. 2002. Ways of meaning, ways of learning: Code as an explanatory concept. *British Journal of Sociology of Education* 23.4: 537–548.

  In this article, Hasan discusses the notion of "code" as a means through which interactions between language, culture, and consciousness determine the language individuals learn. It is argued that although language learning is achieved by "individual minds," these minds are socially fashioned through discursive interaction.

Hoey, M. 2005. *Lexical priming: A new theory of words and language*. London: Routledge. [ISBN: 9780203327630]

  This book introduces a theory of the lexicon that focuses on how words co-occur in common patterns of use. Hoey argues that words are "primed" to occur with certain others, and these primings vary from individual to individual given that each person has unique linguistic experiences and histories that contribute to such primings.

Ludlow, Peter. 2011. *The philosophy of generative linguistics*. Oxford: Oxford Univ. Press. [ISBN: 9780199258536]

This book is a set of reflections on the philosophical issues in generative linguistics. The relevant chapter here is chapter 2—"The Ontology of Generative Linguistics," in which the notion of idiolect is invoked in a contrasting of I-Language and E-Language.

Martin, J. R. 2009. Semantic variation: Modelling realisation, instantiation and individuation in social semiosis. In *New discourse on language: Functional perspectives on multimodality, identity, and affiliation.* Edited by Monika Bednarek and J. R. Martin, 1–34. London: Bloomsbury. [ISBN: 9781441153227]

The introductory chapter to a volume that includes a number of chapters focusing on the notion of "individuation," which is similar to Matthiessen's concept of personalized meaning potential, also from Systemic Functional Linguistics. For Martin, individuation specializes meaning potential—the network of options and resources for making meaning—to *users* rather than *uses* of language.

Matthiessen, Christian M. I. M. 2009. Meaning in the making: Meaning potential emerging from acts of meaning. *Language Learning* 59 (supplement 1): 206–229.

This article is concerned primarily with how individuals learn language, or more specifically, how their "personalized meaning potentials" emerge from "acts of meaning" that make up texts and interaction that they engage with as they go through life.

Schmid, Hans-Jörg, and Annette Mantlik. 2015. Entrenchment in historical corpora? Reconstructing dead authors' minds from their usage profiles. *Anglia* 133.4: 583–623.

This paper compares the usage of N+BE+*that* constructions in the writing of 83 authors from eight historical corpora. Differences between authors in their use of this structure are interpreted in terms of usage-based concepts of entrenchment and conventionalization.

IDIOLECT AND PHILOSOPHY OF LANGUAGE

Idiolect and individuals' language are on occasion discussed within studies in the philosophy of language. As a general introductory reference, the *Stanford Encyclopedia of Philosophy* includes Barber and Garcia 2017, a sizable entry on Idiolects from a philosophy of language perspective. More specifically, Wittgenstein 2001 writes of "private language," comprising words expressing feelings, and experiences of which only the speaker knows and so cannot be understood by anyone else. Generally, philosophical discussions about idiolect hinge on a contrast between internalist and externalist views of language, or view idiolect and the relationship between language and thought in an individual in the context of the communal language and normative rules. Mercier 1993 presents idiolect—or "language individuation"—as a challenge to externalist views of language. Bezuidenhout 2006 discusses and develops Chomsky's distinction between I-Language and E-Language, and explores the connection between internalism and individualism in language and the idiolect as the primary object of scientific study. George 1990 discusses the control that individuals have over their own idiolect, or their knowledge about language, and their ability to alter the properties of their language. Drawing on the notion of "idiolectal error," Barber 2001 deals specifically with ideas of an individual's beliefs and knowledge about language and what can be deduced about their idiolect on the basis of such beliefs. Smith 2001 is a response to Barber 2001 and argues for a more nuanced view of speakers' knowledge of language. Higginbotham 2006 takes an idiolectal but non-internalist view of language, thought, and meaning also through the lens of convention and error, and Briscoe 2006 similarly conceives of a theory of idiolect that disentangles idiolectal meaning from individual understanding. Against a backdrop of major competing views of language in philosophy, Chiffi 2012 poses a series of questions about idiolect vis-à-vis common languages and processes of communication.

Barber, Alex. 2001. Idiolectal error. *Mind & Language* 16.3: 263–283.

An accessible article that teases out the ways in which idiolectal errors, that is, mistakenly held *beliefs* about one's language, can (or cannot) shed light on idiolect as a linguistic theory. The article ultimately argues that there are no such things as "idiolectal error" or "idiolectal accuracy." In the same issue of *Mind & Language,* Smith 2001 offers a response to this essay.

Barber, Alex, and Eduardo Garcia Ramirez. 2017.
Idiolects[https://plato.stanford.edu/archives/spr2017/entries/idiolects]. In *The Stanford encyclopedia of philosophy.* Edited by Edward N. Zalta. Stanford, CA: Stanford Univ.
First published in 2004 but revised in 2017, this encyclopedia entry is a useful introduction to the notion of idiolect in philosophy of language. It provides, too, a good list of sources for readers to follow up.

Bezuidenhout, Anne L. 2006. Language as internal. In *The Oxford handbook of philosophy of language*. Edited by Ernest Lepore and Barry C. Smith, 127–139. Oxford: Oxford Univ. Press. [ISBN: 9780199259410]
One of two chapters in this list from Oxford's *Handbook of Philosophy of Language.* A succinct consolidation and discussion of Chomsky's ideas of internal and external language, which makes a useful link between internalism and idiolect.

Briscoe, Robert Eamon. 2006. Individualism, externalism and idiolectical meaning. *Synthese* 152.1: 95–128.
An article that proposes an approach to idiolectal meaning that can accept "anti-internalist" premises while rejecting "anti-individualist" ones.

Chiffi, Daniele. 2012. Idiolects and language. *Axiomathes* 22.4: 417–432.
A helpful paper to those new to the philosophy of language. Chiffi outlines some of the major competing theories in the field, and provides a good amount of follow-up reading. Written clearly, this is an illuminating discussion of idiolect, accessible to nonphilosophers.

George, Alexander. 1990. Whose language is it anyway? Some notes on idiolects. *Philosophical Quarterly* 40.160: 275–298.
This paper is divided into two sections, each addressing a different view of idiolect and its relationship with language beliefs and control through discussions of errors or malaprops. Concludes with some considerations for the phenomenology of language.

Higginbotham, James. 2006. Languages and idiolects: Their language and ours. In *The Oxford handbook of philosophy of language*. Edited by Ernest Lepore and Barry C. Smith, 140–148. Oxford: Oxford Univ. Press. [ISBN: 9780199259410]

One of two chapters in this list from Oxford's *Handbook of Philosophy of Language*. A short, useful chapter that is accessible for a newcomer to philosophy of language. Taking an anecdote of their son's language use as motivation, Higginbotham discusses a wide range of intersections between the social, the idiolectal, and the conventionalization and conception of language.

Mercier, Adèle. 1993. Normativism and the mental: A problem of language individuation. *Philosophical Studies* 72.1: 71–88.

This paper outlines three externalist arguments of language, before demonstrating how they are challenged by the notion of what Mercier calls "language individuation," and emphasizes the difficulties posed by linguistic creativity of individuals.

Smith, Barry C. 2001. Idiolects and understanding: Comments on Barber. *Mind & Language* 16.3: 284–289.

Best read alongside Barber 2001 in the same issue of *Mind & Language,* Smith offers a response to Barber's main arguments. In particular, Smith reminds readers of the view that individual speakers are authoritative about what words mean and what word strings they find grammatical.

Wittgenstein, Ludwig. 2001. *Philosophical investigations.* 3d ed. Translated by G. E. M. Anscombe. Oxford: Blackwell. [ISBN: 9780631231271]

In this book originally published in 1953 and now in its third edition, Wittgenstein introduces the now famous notion of "private language" in section 243.

IDIOLECT IN CORPUS AND COMPUTATIONAL LINGUISTICS

The advent of corpus linguistics and the ability to collect and analyze far larger datasets of language than previously feasible have, to some degree, reignited interest in the empirical

analyses of idiolect and individuals. Corpus studies of idiolect have mainly focused on demonstrating the distinctiveness of individuals' language, either through the comparison with a small number of other individuals or with a general reference corpus. For the most part, however, the corpora of individuals collected for such studies have been of the idiolects of famous people for whom lots of language data is readily available. Kredens 2002 is one of the earliest in which the idiolects of two famous English musicians, as represented by two small corpora comprising just over 3,000 words each, are compared on the basis of their use of a range of lexico-grammatical features. Mollin 2009 identifies collocation preferences distinctive of Tony Blair's idiolect by comparing the relative frequency of those found in a 3-million-word corpus of his language with their frequency in the British National Corpus. Similarly, Barlow 2013 demonstrates how the idiolects of six U.S. press secretaries can be differentiated on the basis of how frequently they use very common two- and three-word collocations. Coniam 2004 compares a "personal corpus" of the author's own academic writing with corpora of two other academics and the British National Corpus, and shows how an individual's textual profile can emerge from such an analysis. A number of studies have used corpus data as a means through which to investigate individuals' mental representation, storage, and processing of language, such as the focus on formulaic sequences in Schmitt, et al. 2004, high-frequency collocations in Durrant and Doherty 2010, and morphological derivations in De Smet 2016. Computational linguistics has picked up from corpus linguistics and has moved from distinguishing between idiolects to identifying the individuals responsible for writing anonymous or disputed documents—a task known as authorship attribution. "Stylometry" is the umbrella term for the statistical analysis and identification of idiolects. The seminal paper in this regard is the Mosteller and Wallace 1964 analysis of *The Federalist Papers*. El Bouanani and Kassou 2014 offers a useful survey of modern-day authorship analysis techniques that have developed in the decades since.

Barlow, Michael. 2013. Individual differences and usage-based grammar. *International Journal of Corpus Linguistics* 18.4: 443–478.

A corpus-based study of idiolect in which individuals are compared and distinguished based on their use of very frequent two- and three-word utterances. As well as being methodologically important, the approach to idiolect in this study is framed in terms of usage-based grammar, and particularly exemplar theory.

Coniam, David. 2004. Concordancing oneself: Constructing individual textual profiles. *International Journal of Corpus Linguistics* 9.2: 271–298.

An interesting study in that the main corpus under analysis is one comprising the author's own academic writing, which is compared with two other academics and a sample of the written component of the British National Corpus. The linguistic features drawn on to compare idiolects include keywords, *n*-gram clusters, pronouns, and hedges.

De Smet, Hendrik. 2016. The root of ruthless: Individual variation as a window on mental representation. *International Journal of Corpus Linguistics* 21.2: 250–271.

This study investigates whether the mental representations of different types of derivational morphology are related to each other in the idiolects of 174 individuals from a corpus of *New York Times* writing. It is a good example of using a corpus approach to analyze and compare inter- and intra-idiolectal variation.

Durrant, Phillip, and Alice Doherty. 2010. Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory* 6.2: 125–155.

This paper empirically tests the relationship between the frequency of collocations in corpus data and the storage of these collocations in the minds of individual speakers. The study is an example of a successful triangulation of corpus-based and experimental methods in analyzing the language of individuals.

El Bouanani, Sara El Manar, and Ismail Kassou. 2014. Authorship analysis studies: A survey. *International Journal of Computer Applications* 86.12: 22–29.

Over the years, many surveys of stylometric techniques have been published. This one is the most comprehensive and recent. It details the linguistic features commonly used in authorship studies and the statistical procedures applied. It is full of useful references, and includes a summary of some of the most important works in this area.

Kredens, Krzysztof. 2002. Towards a corpus-based methodology of forensic authorship
attribution: A comparative study of two idiolects. In *PALC'01:International practical
applications in language corpora,* Lodz, Poland, 7–9 September 2001. Edited by Barbara
Lewandowska-Tomaszczyk, 405–437. Peter Lang: Frankfurt am Mein. [ISBN:
9788371714887]
This study is framed from the perspective of forensic linguistics and is an early example of the
systematic corpus-based comparison of two idiolects across a range of linguistic variables. As
well as being pioneering methodologically, it provides important discussion of the implications
of its findings for application in forensic contexts.

Mollin, Sandra. 2009. "I entirely understand" is a Blairism: The methodology of identifying
idiolectal collocations. *International Journal of Corpus Linguistics* 14.3: 367–392.
This paper follows a systematic, replicable, and robust methodological procedure for
identifying collocational preferences that are idiolectally distinctive of an individual. It
demonstrates the value of undertaking controlled comparisons between idiolect-specific
corpora and a general reference corpus.

Mosteller, Frederick, and David L. Wallace. 1964. *Inference and disputed authorship: The
Federalist*. Reading, MA: Addison-Wesley.
The object of analysis in this study is an important set of documents in U.S history. Carried out
by statisticians, rather than linguists, the approach and findings of this study laid the
foundations for modern-day stylometric methods. This involves both the statistical
measurement of similarity and distance between idiolects, and the attribution of disputed
documents based on this known variation between the idiolects of potential authors.

Schmitt, Norbert, Sarah Grandage, and Svenja Adolphs. 2004. Are corpus-derived recurrent
clusters psycholinguistically valid? In *Formulaic sequences: Acquisition, processing, and use.*
Edited by Norbert Schmitt, 127–151. Amsterdam: John Benjamins. [ISBN: 9789027295750]
This chapter investigates whether corpus-derived recurrent lexical strings are stored holistically
in speakers' minds as formulaic sequences. The combination of corpus linguistic data and
psycholinguistic experiments produces important results for the notion of idiolect. It is argued

that particular word clusters may be holistically stored for some people but not others, and that individuals have their own "formulalect" or "phrasalect."

IDIOLECT IN FORENSIC LINGUISTICS

An area of linguistics in which there is a sustained theoretical debate on idiolect is forensic linguistics, and particularly forensic authorship analysis. Coulthard 2004 is an important work in this regard, drawing on famous criminal cases and instances of student plagiarism to demonstrate the value that a theory of idiolect has in underpinning forensic linguistic investigations. Turell 2010 is another forensic linguistic article that deals directly with the concept idiolect, introducing a new concept of "idiolectal style" as a less abstract, more realistically observable and measurable manifestation of an individual's language. Grant 2010 offers a different view, as it is argued that forensic authorship analysis need not rely on a theory of idiolect to be successful, but rather can focus only on the consistency and distinctiveness of the authors' styles in any given case. In later work on the policing of online pedophile activity and identity disguise, Grant and MacLeod 2018 develop a 'resource constraint model of language and identity' in which an individual's linguistic style and identity are performed through a process of them drawing on a set of resources available in interaction, including their sociolinguistic history, their physical self and the context of a given communicative event. McMenamin 2002 is a major text in "forensic stylistics" and situates the identification of authors' style markers within linguistic variation, and as underpinned by the theory of idiolect. Johnson and Wright 2014 undertakes a case study of one author in the Enron email corpus, and proposes that word *n*-grams can capture distinctive "chunks" of individual idiolects that reflect their unique sociolinguistic histories and can, in turn, be used to identify authors of anonymous texts in forensic casework. Wright 2017 develops this notion further by drawing on usage-based theories of grammar, and entrenchment in particular, to account for idiolect-distinctive word strings that show different individuals encoding the same meaning in different ways. Idiolect features, too, in forensic phonetics, and more specifically, in speaker comparison. Baldwin 1979 is an early work that directly addresses the notion of idiolect in speaker recognition. Watt 2009 offers a more recent discussion of identifying individuals through speech (though exercises caution with the term "idiolect"), and experimental studies such as Gavaldà 2016 test the discriminatory potential of specific speech features.

Baldwin, J. R. 1979. Phonetics and speaker identification. *Medicine, Science and the Law* 19.4: 231–232.

A short article and an early exploration of phonetics and forensic applications of speaker identification. The article briefly explains, using some case examples, how elements of an individual's idiolect can be used to identify them in certain circumstances.

Coulthard, Malcolm. 2004. Author identification, idiolect, and linguistic uniqueness. *Applied Linguistics* 24.4: 431–447.

Both methodologically and theoretically, this is an important paper for forensic authorship analysis. After demonstrating the potentially author-unique nature of word strings and their usefulness in author identification through the discussion of high-profile cases, Coulthard specifically addresses issues of reliability and admissibility of forensic linguistic methods and evidence.

Gavaldà, Núria. 2016. Individual variation in allophonic processes of /t/ in standard southern British English. *International Journal of Speech, Language and the Law* 23.1: 43–69.

An experimental analysis that demonstrates the ways in which the pronunciation of /t/ can distinguish between the idiolects of ten individual speakers in certain contexts. This is a recent example of the applicability of sociolinguistic variation analysis to discriminate between the speech of individuals in a forensic context.

Grant, Tim. 2010. Txt 4n6: Idiolect free authorship analysis? In *The Routledge handbook of forensic linguistics.* Edited by Malcolm Coulthard and Alison Johnson, 508–522. London: Routledge. [ISBN: 9780415463096]

This chapter challenges the assumption that a theory of idiolect is required for successful authorship analysis. After discussing different perspectives on idiolect, Grant argues that measurable notions of consistency and distinctiveness of styles, without recourse to a theory of "idiolect" per se, are sufficient in forensic case work. An approach is then demonstrated using SMS text-message data from a famous murder case in the UK.

Grant, Tim, and Nicci MacLeod. 2018. Resources and constraints in linguistic identity performance: a theory of authorship. *Language and Law/Linguagem e Direito* 5.1: 80–96.
This paper distinguishes between models of language and identity commonly drawn upon in sociolinguistics from those found in discourse analysis. Bridging the gap between so-called "essentialist" and "interactionist" views respectively, this article develops a theory of identity performance as interactionally emergent, relying on the resources drawn upon by speakers. This theory is then applied to the context of undercover police officers performing deceptive linguistic identities when investigating online child sex abusers.

Johnson, A., and D. Wright. 2014. Identifying idiolect in forensic authorship attribution: An *n*-gram textbite approach. *Language and Law/Linguagem e Direito* 1.1: 37–69.
One in a series of studies that use the publicly available Enron corpus as experimental data for analyzing idiolect. Central to this paper is the in-depth analysis of the idiolect of one individual in the corpus, and in particular the uniformity and distinctiveness of their professional email style. Implications of findings are discussed in relation to forensic authorship analysis

McMenamin, Gerald R. 2002. *Forensic linguistics: Advances in forensic stylistics*. Boca Raton, FL: CRC Press. [ISBN: 9780849309663]
An important work in the field of forensic stylistics and authorship analysis. The methods laid out in this book have been widely taken up in research and casework since, and are normally compared and contrasted with the other major methodological approach to authorship and idiolect analysis—stylometry.

Turell, M. Teresa. 2010. The use of textual, grammatical, and sociolinguistic evidence in forensic text comparison. *International Journal of Speech, Language and the Law* 17.2: 211–250.
This important paper problematizes the abstract nature of the theory of idiolect and its applicability in a forensic context. After introducing "idiolectal style" as an observable and measurable specification of idiolect, Turell reports the innovative methods used and the subsequent results of a forensic authorship case in Spain.

Watt, Dominic. 2009. The identification of the individual through speech. In *Language and Identities*. Edited by Carmen Llamas and Dominic Watt, 76–85. Edinburgh: Edinburgh Univ. Press. [ISBN: 9780748635764]

This chapter provides a detailed and measured discussion of speaker identification. An important distinction is made between "naive" speaker identification by nonexperts, and methods of technical speaker identification undertaken by expert phoneticians and applied in forensic context. A section is also dedicated to the description of sources of individual variation in speech and voice.

Wright, David. 2017. Using word *n*-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem. *International Journal of Corpus Linguistics* 22.2: 212–241. Another article that uses the Enron corpus as experimental data. Developing the method and theory first proposed by Johnson and Wright 2014, this paper demonstrates how idiolectally distinctive word strings can identify the authors of anonymized email samples. A subsequent corpus analysis reveals the specific differences across authors in fulfilling the same speech act, and interprets the findings in relation to theories of entrenchment, formulaic language, and lexical priming.